The Baillon-Haddad Theorem Revisited

Heinz H. Bauschke[*]
Mathematics, University of British Columbia Okanagan
Kelowna, B.C. V1V 1V7, Canada
heinz.bauschke@ubc.ca

Patrick L. Combettes[†]
UPMC Université Paris 06
Laboratoire Jacques-Louis Lions – UMR CNRS 7598
75005 Paris, France
plc@math.jussieu.fr

Abstract

In 1977, Baillon and Haddad proved that if the gradient of a convex and continuously differentiable function is nonexpansive, then it is actually firmly nonexpansive. This result, which has become known as the Baillon-Haddad theorem, has found many applications in optimization and numerical functional analysis. In this note, we propose short alternative proofs of this result and strengthen its conclusion.

2000 Mathematics Subject Classification:
Primary 47H09, 90C25; Secondary 26A51, 26B25, 46C05, 47H05, 52A41.

Key Words:
Backward-backward splitting, Bregman distance, cocoercivity, convex function, Dunn property, firmly nonexpansive, forward-backward splitting, gradient, inverse strongly monotone, Moreau envelope, nonexpansive, proximal mapping, proximity operator.

[*]Research partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program.

[†]Research supported by the Agence Nationale de la Recherche under grant ANR-08-BLAN-0294-02.

1 Introduction

Throughout, $\mathcal{H}$ is a real Hilbert space with scalar product $\langle \cdot \mid \cdot \rangle$ and induced norm $\|\cdot\|$. Let C be a nonempty subset of $\mathcal{H}$, let $T\colon C \to \mathcal{H}$, and let $\beta \in \left]0,+\infty\right[$. Then T is $1/\beta$-cocoercive if (this property is also known as the Dunn property or inverse strong monotonicity)

$$(\forall x \in C)(\forall y \in C) \quad \beta \langle x-y \mid Tx-Tx\rangle \geq \|Tx-Ty\|^2, \tag{1}$$

and T is β-Lipschitz continuous if

$$(\forall x \in C)(\forall y \in C) \quad \|Tx-Tx\| \leq \beta\|x-y\|^2. \tag{2}$$

When $\beta = 1$, (1) means that T is firmly nonexpansive and (2) that T is nonexpansive. Cocoercivity arises in various areas of optimization and nonlinear analysis, e.g., [2, 5, 6, 9, 14, 15, 20, 23]. It follows from the Cauchy-Schwarz inequality that $1/\beta$-cocoercivity implies β-Lipschitz continuity. However, the converse fails; take for instance $T = -\mathrm{Id}$, which is nonexpansive but not firmly nonexpansive. In 1977, Baillon and Haddad showed that, if $C = \mathcal{H}$ and T is the gradient of a convex function, then (1) and (2) coincide. This remarkable result, which has important applications in optimization (see for instance [7, 21]), has become known as the Baillon-Haddad theorem.

Theorem 1.1 (Baillon-Haddad) [3, Corollaire 10] *Let $f\colon \mathcal{H} \to \mathbb{R}$ be convex, Fréchet differentiable on $\mathcal{H}$, and such that ∇f is β-Lipschitz continuous for some $\beta \in \left]0,+\infty\right[$. Then ∇f is $1/\beta$-cocoercive.*

In [3], Theorem 1.1 was derived from a more general result concerning n-cyclically monotone operators in normed vector spaces. Since then, direct proofs have been proposed, such as [11, Lemma 6.7], [12, Theorem X.4.2.2], and [18, Proposition 12.60] for Euclidean spaces. These approaches rely on convex analytical and integration arguments. An infinite dimensional proof can be found in [22, Remark 3.5.2], as a corollary to results on the properties of uniformly smooth convex functions.

The goal of our paper is to provide new insights into the Baillon-Haddad theorem. In Section 2, we propose a short new proof of Theorem 1.1 and present additional equivalent conditions, thus making a connection with lesser known parts of Moreau's classical paper [16]. In Section 3, we provide a second order variant of the Baillon-Haddad theorem that partially extends work by Dunn [9].

Notation and background. Our notation is standard: $\Gamma_0(\mathcal{H})$ is the class of proper lower semicontinuous convex functions from $\mathcal{H}$ to $\left]-\infty,+\infty\right]$ and $\square$ denotes infimal convolution. The conjugate of a function $f\colon \mathcal{H} \to \left]-\infty,+\infty\right]$ is denoted by f^*, and its subdifferential by ∂f. For background on convex analysis, we refer the reader to [12, 17, 22].

2 An enhanced Baillon-Haddad theorem

Let us start with some standard facts on Moreau envelopes and proximity operators; we refer the reader to Moreau's original paper [16] and to [1, 7, 18] for details and complements. Let $\varphi \in \Gamma_0(\mathcal{H})$

and let $\gamma \in \,]0,+\infty[$. The Moreau envelope of φ of index γ is the finite continuous convex function

$$\operatorname{env}_\gamma(\varphi) = \varphi \,\square\, \big(q/\gamma\big), \quad \text{where} \quad q = \frac{1}{2}\|\cdot\|^2. \tag{3}$$

Moreau's decomposition asserts that

$$\operatorname{env}_{1/\gamma}(\varphi) + \operatorname{env}_\gamma(\varphi^*) \circ (\gamma \operatorname{Id}) = \gamma q. \tag{4}$$

The proximity operator (or proximal mapping) of f is the operator $\operatorname{Prox}_\varphi = (\operatorname{Id} + \partial\varphi)^{-1}$; it maps each $x \in \mathcal{H}$ to the unique minimizer of the function $y \mapsto \varphi(y) + q(x-y)$. The Moreau envelope $\operatorname{env}_1(\varphi)$ is Fréchet differentiable with gradient $\nabla \operatorname{env}_1(\varphi) = \operatorname{Prox}_{\varphi^*}$. Hence, (4) yields

$$\nabla \operatorname{env}_{1/\gamma}(\varphi) = \operatorname{Prox}_{\gamma\varphi^*} \circ (\gamma \operatorname{Id}) = \gamma(\operatorname{Id} - \operatorname{Prox}_{\varphi/\gamma}). \tag{5}$$

Moreover,

$$\operatorname{Prox}_\varphi \colon \mathcal{H} \to \mathcal{H} \ \text{ is firmly nonexpansive.} \tag{6}$$

We are now ready to present the main result of this section, which strengthens the conclusion of Theorem 1.1 by providing four additional equivalent conditions and a short new proof.

Theorem 2.1 *Let $f \in \Gamma_0(\mathcal{H})$, let $\beta \in \,]0,+\infty[$, and set $h = f^* - q/\beta$. Then the following are equivalent.*

(i) *f is Fréchet differentiable on $\mathcal{H}$ and ∇f is β-Lipschitz continuous.*

(ii) *$\beta q - f$ is convex.*

(iii) *$f^* - q/\beta$ is convex (i.e., f^* is $1/\beta$-strongly convex).*

(iv) *$h \in \Gamma_0(\mathcal{H})$ and $f = \operatorname{env}_{1/\beta}(h^*) = \beta q - \operatorname{env}_\beta(h) \circ \beta \operatorname{Id}$.*

(v) *$h \in \Gamma_0(\mathcal{H})$ and $\nabla f = \operatorname{Prox}_{\beta h} \circ \beta \operatorname{Id} = \beta(\operatorname{Id} - \operatorname{Prox}_{h^*/\beta})$.*

(vi) *f is Fréchet differentiable on $\mathcal{H}$ and ∇f is $1/\beta$-cocoercive.*

Proof. (i)⇒(ii): By Cauchy-Schwarz, $(\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ \langle x-y \mid \beta x - \nabla f(x) - \beta y + \nabla f(y)\rangle = \beta\|x-y\|^2 - \langle x-y \mid \nabla f(x) - \nabla f(y)\rangle \geq \|x-y\|(\beta\|x-y\| - \|\nabla f(x) - \nabla f(y)\|) \geq 0$. Hence, $\nabla(\beta q - f) = \beta \operatorname{Id} - \nabla f$ is monotone and it follows that $\beta q - f$ is convex (see, e.g., [22, Theorem 2.1.11]).

(ii)⇒(iii): Set $g = \beta q - f$. Then $g \in \Gamma_0(\mathcal{H})$ and therefore $g = g^{**}$. Accordingly,

$$f = \beta q - g = \beta q - g^{**} = \beta q - \sup_{u\in\mathcal{H}} \big(\langle \cdot \mid u\rangle - g^*(u)\big) = \inf_{u\in\mathcal{H}} \big(\beta q - \langle \cdot \mid u\rangle + g^*(u)\big). \tag{7}$$

Hence

$$\begin{aligned} f^* &= \sup_{u\in\mathcal{H}} \big(\beta q - \langle \cdot \mid u\rangle + g^*(u)\big)^* = \sup_{u\in\mathcal{H}} \big(\big(\beta q - \langle \cdot \mid u\rangle\big)^* - g^*(u)\big) \\ &= \sup_{u\in\mathcal{H}} \big(q(\cdot + u)/\beta - g^*(u)\big) = q/\beta + \sup_{u\in\mathcal{H}} \big(\big(\langle \cdot \mid u\rangle + q(u)\big)/\beta - g^*(u)\big), \end{aligned} \tag{8}$$

where the last term is convex as a supremum of affine functions. Thus, h is convex.

(iii)⇒(iv): Since $f \in \Gamma_0(\mathcal{H})$ and h is convex, we have $h \in \Gamma_0(\mathcal{H})$, $h^* \in \Gamma_0(\mathcal{H})$, and $f = f^{**} = (h + q/\beta)^* = h^* \,\square\, \beta q = \operatorname{env}_{1/\beta}(h^*) = \beta q - \operatorname{env}_\beta(h) \circ \beta \operatorname{Id}$, where the last identity follows from (4).

(iv)⇒(v): Use (5).

(v)⇒(vi): By (6), $\operatorname{Prox}_{\beta h}$ is firmly nonexpansive. Hence, it follows from (1) that $\nabla f = \operatorname{Prox}_{\beta h} \circ \beta \operatorname{Id}$ is $1/\beta$-cocoercive.

(vi)⇒(i): Apply the Cauchy-Schwarz inequality. ■

Remark 2.2 Some comments regarding Theorem 2.1 are in order.

(a) The proof of the implication (i)⇒(vi), i.e., of the Baillon-Haddad theorem (Theorem 1.1) appears to be new and shorter than those found in the literature. In addition, Theorem 2.1 brings to light various characterizations of the Lipschitz continuity of the gradient of a convex function. The equivalences (ii)⇔(iii)⇔(iv) are due to Moreau, who established them (for $\beta = 1$) in [16, Proposition 9.b] (see also [13, Corollary 3]). On the other hand, the equivalences (i)⇔(iii)⇔(iv)⇔(vi) are shown in Euclidean spaces in [18, Proposition 12.60] with different techniques.

(b) Set $\beta = 1$. The conclusion of Theorem 1.1 is that $\nabla f : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive. Hence, since the class of firmly nonexpansive operators with domain $\mathcal{H}$ coincides with that of resolvents of maximal monotone operators [10, Section 1.11], we have $\nabla f = (\operatorname{Id} + A)^{-1}$, for some maximal monotone operator $A : \mathcal{H} \to 2^{\mathcal{H}}$. However, (v) more precisely reveals ∇f to be the proximity operator of h, i.e., $A = \partial h = \partial f^* - \operatorname{Id}$.

(c) Let $f_1 \in \Gamma_0(\mathcal{H})$, let $f_2 : \mathcal{H} \to \mathbb{R}$ be convex and differentiable with a Lipschitz continuous gradient, and consider the problem of minimizing $f_1 + f_2$. Without loss of generality (rescale), we assume that the Lipschitz constant of ∇f_2 is $\beta = 1$. A standard algorithm for solving this problem is the forward-backward algorithm [7, 20]

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}_{\gamma_n f_1}\big(x_n - \gamma_n \nabla f_2(x_n)\big), \quad 0 < \gamma_n < 2. \tag{9}$$

Now set $h_2 = f_2^* - q$. Then it follows from the implication (i)⇒(v) that $\nabla f_2 = \operatorname{Id} - \operatorname{Prox}_{h_2^*}$. Hence, we can rewrite (9) as

$$x_0 \in \mathcal{H} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = \operatorname{Prox}_{\gamma_n f_1}\big((1 - \gamma_n)x_n + \gamma_n \operatorname{Prox}_{h_2^*} x_n\big), \quad 0 < \gamma_n < 2. \tag{10}$$

This shows that the forward-backward algorithm (9) is actually a backward-backward algorithm. In particular, for $\gamma_n \equiv 1$, we recover the basic backward-backward iteration $x_{n+1} = \operatorname{Prox}_{f_1} \operatorname{Prox}_{h_2^*} x_n$.

We conclude this section with an alternative formulation of the Baillon-Haddad theorem that brings into play Bregman distances. Recall that if $\varphi \in \Gamma_0(\mathcal{H})$ is Gâteaux differentiable on

$\operatorname{int}\operatorname{dom}\varphi \neq \varnothing$, the associated Bregman distance D_φ is defined by

$$D_\varphi \colon \mathcal{H}\times\mathcal{H} \to [0,+\infty] : (x,y) \mapsto \begin{cases} \varphi(x) - \varphi(y) - \langle x-y \mid \nabla\varphi(y)\rangle, & \text{if } y \in \operatorname{int}\operatorname{dom}\varphi;\\ +\infty, & \text{otherwise.}\end{cases} \tag{11}$$

Corollary 2.3 *Let $\beta \in \left]0,+\infty\right[$, and let $f\colon \mathcal{H} \to \mathbb{R}$ be convex, Fréchet differentiable on $\mathcal{H}$, and such that f^* is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f^* \neq \varnothing$. Then the following are equivalent.*

(i) *∇f is β-Lipschitz continuous.*

(ii) *$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ D_f(x,y) \leq \beta q(x-y)$.*

(iii) *$(\forall x^* \in \mathcal{H})(\forall y^* \in \mathcal{H})\ \beta D_{f^*}(x^*,y^*) \geq q(x^*-y^*)$.*

Proof. (i)⇔(ii): Set $g = \beta q - f$. Then g is Fréchet differentiable on $\operatorname{dom} g = \mathcal{H}$ and $\nabla g = \beta\,\mathrm{Id} - \nabla f$. Hence, it follows from the equivalence (i)⇔(ii) in Theorem 2.1 and (11) that (i) $\Leftrightarrow$ $g \in \Gamma_0(\mathcal{H})$ is Fréchet differentiable on $\operatorname{int}\operatorname{dom} f = \mathcal{H} \Leftrightarrow (\forall x \in \mathcal{H})(\forall y \in \mathcal{H})\ D_g(x,y) \geq 0 \Leftrightarrow (\forall x \in \mathcal{H})(\forall y \in \mathcal{H})$ $D_f(x,y) \leq \beta q(x-y)$.

(i)⇔(iii): Set $h = f^* - q/\beta$. Then h is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} h = \operatorname{int}\operatorname{dom} f^*$, with $\nabla h = \nabla f^* - (1/\beta)\,\mathrm{Id}$. Hence, in view of the equivalence (i)⇔(iii) in Theorem 2.1 and (11), (i) $\Leftrightarrow$ $h \in \Gamma_0(\mathcal{H})$ is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} h = \operatorname{int}\operatorname{dom} f^* \Leftrightarrow (\forall x^* \in \mathcal{H})(\forall y^* \in \mathcal{H})$ $D_h(x^*,y^*) \geq 0 \Leftrightarrow (\forall x^* \in \mathcal{H})(\forall y^* \in \mathcal{H})\ D_{f^*}(x^*,y^*) \geq q(x^*-y^*)/\beta$. ■

3 A second order Baillon-Haddad theorem

Under the more restrictive assumption that the underlying convex function is twice continuously differentiable, we shall obtain in Theorem 3.3 a very short and transparent proof inspired by the work of Dunn [9]. We require two preliminary propositions.

Proposition 3.1 *Let C be a nonempty open convex subset of $\mathcal{H}$, let $\mathcal{B}$ be a real Banach space, and let $G\colon C \to \mathcal{B}$ be continuously Fréchet differentiable on C. Then G is nonexpansive if and only if $(\forall x \in C)\ \|\nabla G(x)\| \leq 1$.*

Proof. Let $x \in C$ and let $y \in \mathcal{H}$. Suppose that G is nonexpansive. For every $t \in \left]0,+\infty\right[$ sufficiently small, $x+ty \in C$ and hence $\|G(x+ty)-G(x)\|/t \leq \|y\|$. Letting $t \downarrow 0$, we deduce that $\|(\nabla G(x))y\| \leq \|y\|$. Since y was chosen arbitrarily, we conclude that $\|\nabla G(x)\| \leq 1$. Conversely, if $y \in C$, we derive from the mean value theorem (see, e.g., [8, Theorem 5.1.12]) that $\|G(y)-G(x)\| \leq \|y-x\| \sup_{z\in[x,y]} \|\nabla G(z)\| \leq \|y-x\|$. ■

Let $A\colon \mathcal{H} \to \mathcal{H}$ and $B\colon \mathcal{H} \to \mathcal{H}$ be self-adjoint bounded linear operators. Then A is positive, written $A \succeq 0$, if $(\forall x \in \mathcal{H})\ \langle x \mid Ax\rangle \geq 0$. We write $A \succeq B$ if $A - B \succeq 0$. The following result is part of the folklore.

Proposition 3.2 *Let $A\colon \mathcal{H} \to \mathcal{H}$ be a bounded self-adjoint linear operator. Then $\|A\| \leq 1$ if and only if $\mathrm{Id} \succeq A \succeq -\mathrm{Id}$.*

Proof. Assume that $\mathcal{H} \neq \{0\}$ and set $S = \big\{x \in \mathcal{H} \mid \|x\| = 1\big\}$. Then $\mathrm{Id} \succeq A \Leftrightarrow (\forall x \in \mathcal{H})\ \langle x \mid x\rangle \geq \langle x \mid Ax\rangle \Leftrightarrow (\forall x \in S)\ 1 = \langle x \mid x\rangle \geq \langle x \mid Ax\rangle$. Similarly, $A \succeq -\mathrm{Id} \Leftrightarrow (\forall x \in S)\ \langle x \mid Ax\rangle \geq -1$. Hence $\mathrm{Id} \succeq A \succeq -\mathrm{Id} \Leftrightarrow (\forall x \in S)\ |\langle x \mid Ax\rangle| \leq 1 \Leftrightarrow \|A\| = \sup_{x\in S} |\langle x \mid Ax\rangle| \leq 1$. ■

The main result of this section is a Baillon-Haddad theorem for twice continuously Fréchet differentiable convex functions. It extends [9, Theorem 4], which assumed in addition that f has full domain and uniformly bounded Hessians.

Theorem 3.3 *Let C be a nonempty open convex subset of $\mathcal{H}$, let $f\colon C \to \mathbb{R}$ be convex and twice continuously Fréchet differentiable on C, and let $\beta \in \left]0, +\infty\right[$. Then ∇f is β-Lipschitz continuous if and only if it is $1/\beta$-cocoercive.*

Proof. Define two operators on C by $G = (1/\beta)\nabla f$ and by $H = \nabla G = (1/\beta)\nabla^2 f$. Under our assumptions, the convexity of f is characterized by [22, Theorem 2.1.11]

$$(\forall x \in \mathcal{H}) \quad H(x) \succeq 0. \tag{12}$$

Hence,

$$\begin{aligned}
\nabla f \text{ is } \beta\text{-Lipschitz continuous} &\Leftrightarrow G \text{ is nonexpansive} \\
&\Leftrightarrow (\forall x \in C) \quad \|H(x)\| \leq 1 && \text{(by Proposition 3.1)} \\
&\Leftrightarrow (\forall x \in C) \quad -\mathrm{Id} \preceq H(x) \preceq \mathrm{Id} && \text{(by Proposition 3.2)} \\
&\Leftrightarrow (\forall x \in C) \quad 0 \preceq H(x) \preceq \mathrm{Id} && \text{(by (12))} \\
&\Leftrightarrow (\forall x \in C) \quad -\mathrm{Id} \preceq 2H(x) - \mathrm{Id} \preceq \mathrm{Id} \\
&\Leftrightarrow (\forall x \in C) \quad \|2H(x) - \mathrm{Id}\,\| \leq 1 && \text{(by Proposition 3.2)} \\
&\Leftrightarrow 2G - \mathrm{Id} \text{ is nonexpansive} && \text{(by Proposition 3.1)} \\
&\Leftrightarrow G \text{ is firmly nonexpansive} && \text{(by [10, Lemma 1.11.1])} \\
&\Leftrightarrow \nabla f \text{ is } 1/\beta\text{-cocoercive,} && \text{(by (1))}
\end{aligned}$$

and we obtain the conclusion. ■

In linear functional analysis, the following property is usually obtained via spectral theory.

Corollary 3.4 *Let $A\colon \mathcal{H} \to \mathcal{H}$ be a positive self-adjoint bounded linear operator. Then $(\forall x \in \mathcal{H})$ $\|A\|\ \langle x \mid Ax\rangle \geq \|Ax\|^2$.*

Proof. This is an application of Theorem 3.3 with $f\colon \mathcal{H} \to \mathbb{R}\colon x \mapsto \langle x \mid Ax\rangle/2$. Indeed, f is twice continuously Fréchet differentiable on $\mathcal{H}$ with $\nabla f = A$, which is $\|A\|$-Lipschitz continuous. ■

Remark 3.5 It would be interesting to see whether Theorem 3.3 holds true when the second-order assumption is replaced by Fréchet differentiability. However, the natural approach by approximation does not appear to be applicable; see [4, Section 5] for pertinent comments.

References

[1] H. Attouch: Variational Convergence for Functions and Operators, Pitman, Boston, MA (1984).

[2] H. Attouch, L. M. Briceño-Arias, and P. L. Combettes: A parallel splitting method for coupled monotone inclusions, SIAM J. Control Optim. to appear.

[3] J.-B. Baillon and G. Haddad: Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones, Israel J. Math. 26 (1977) 137–150.

[4] J. M. Borwein and D. Noll: Second order differentiability of convex functions in Banach spaces, Trans. Amer. Math. Soc. 342 (1994) 43–81.

[5] C. L. Byrne: Applied Iterative Methods, A. K. Peters, Wellesley, MA (2008).

[6] P. L. Combettes: Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53 (2004) 475–504.

[7] P. L. Combettes and V. R. Wajs: Signal recovery by proximal forward-backward splitting, Multiscale Model. Simul. 4 (2005) 1168–1200.

[8] Z. Denkowski, S. Migorski, and N. S. Papageorgiou: An Introduction to Nonlinear Analysis: Theory, Kluwer Academic Publishers, Boston, MA (2003).

[9] J. C. Dunn: Convexity, monotonicity, and gradient processes in Hilbert space, J. Math. Anal. Appl. 53 (1976) 145–158.

[10] K. Goebel and S. Reich: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings, Marcel Dekker, New York (1984).

[11] E. G. Golshtein and N. V. Tretyakov: Modified Lagrangians and Monotone Maps in Optimization, Wiley-Interscience, New York (1996).

[12] J.-B. Hiriart-Urruty and C. Lemaréchal: Convex Analysis and Minimization Algorithms, Springer-Verlag, New York (1993).

[13] J.-B. Hiriart-Urruty and Ph. Plazanet: Moreau's decomposition theorem revisited, Ann. Inst. H. Poincaré Anal. Non Linéaire 6 (1989) 325–338.

[14] P.-L. Lions and B. Mercier: Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.

[15] F. Liu and M. Z. Nashed: Regularization of nonlinear ill-posed variational inequalities and convergence rates, Set-Valued Anal. 6 (1998) 313–344.

[16] J.-J. Moreau: Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.

[17] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton, New Jersey (1970).

[18] R. T. Rockafellar and R. J-B Wets: Variational Analysis, 3rd printing, Springer-Verlag, Berlin (2009).

[19] S. Simons: From Hahn-Banach to Monotonicity, Lecture Notes in Math. 1693, Springer-Verlag, New York (2008).

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

[21] I. Yamada and N. Ogura: Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings, Numer. Funct. Anal. Optim. 25 (2004) 619–655.

[22] C. Zălinescu, Convex Analysis in General Vector Spaces, World Scientific Publishing, River Edge, NJ (2002).

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