Lower Bounds on the Haraux Function*

Patrick L. Combettes[1] and Julien N. Mayrand[2]

[1]North Carolina State University
Department of Mathematics
Raleigh, NC 27695, USA
plc@math.ncsu.edu

[2]North Carolina State University
Department of Mathematics
Raleigh, NC 27695, USA
jnmayran@ncsu.edu

Abstract The Haraux function is an important tool in monotone operator theory and its applications. One of its salient properties for a maximally monotone operator is to be valued in $[0,+\infty]$ and to vanish only on the graph of the operator. Sharper lower bounds for this function have been proposed in specific cases. We derive lower bounds in the general context of set-valued operators in reflexive real Banach spaces. These bounds are new, even for maximally monotone operators acting on Euclidean spaces, a scenario in which we show that they can be better than existing ones. As a by-product, we obtain lower bounds on the Fenchel–Young function in variational analysis. Several examples are given and applications to composite monotone inclusions are discussed.

Keywords. Fenchel–Young inequality, Fitzpatrick function, Haraux function, metric resolvent, monotone operator, warped resolvent.

*Contact author: P. L. Combettes. Email: plc@math.ncsu.edu. Phone: +1 919 515 2671. This work was supported by the National Science Foundation under grant DMS-2513409.

§1. Introduction

Throughout, $\mathcal{X} \neq \{0\}$ is a reflexive real Banach space with topological dual $\mathcal{X}^*$ and $\langle \cdot, \cdot \rangle$ denotes the canonical duality pairing. The power set of $\mathcal{X}^*$ is denoted by $2^{\mathcal{X}^*}$. See Section 2 for further notation.

Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be an operator and let $\operatorname{gra} A = \{(x, x^*) \in \mathcal{X} \times \mathcal{X}^* \mid x^* \in Ax\}$ be its graph. The *Haraux function* associated with A is

$$H_A\colon \mathcal{X} \times \mathcal{X}^* \to [-\infty, +\infty]\colon (x, u^*) \mapsto \sup_{(y,y^*)\in\operatorname{gra} A} \langle x - y, y^* - u^* \rangle. \tag{1.1}$$

This function was originally conceived by Haraux in 1974 in unpublished notes as a tool to generalize the notions of 3 monotone and angle bounded operators [27]. It is at the core of many important developments in the theory of monotone operators in reflexive Banach spaces [3, 9, 14, 19, 22, 32, 36, 38, 41]. In particular, it is an essential component of the famous Brézis–Haraux theorem [13], which studies how close the range of the sum of two monotone operators in a Hilbert space is to the Minkowski sum of their ranges. A closely related function, later introduced independently in [26], is the *Fitzpatrick function*

$$F_A = H_A + \langle \cdot, \cdot \rangle. \tag{1.2}$$

A basic property of the Haraux function is that, if A is maximally monotone, then

$$H_A \geqslant 0 \quad \text{and} \quad \operatorname{gra} A = \{(x, u^*) \in \mathcal{X} \times \mathcal{X}^* \mid H_A(x, u^*) = 0\}. \tag{1.3}$$

This follows for instance from [26, Corollary 3.9] and (1.2). A natural question is whether the inequality $H_A \geqslant 0$ can be improved. A related question concerns the Fenchel–Young inequality. Consider a proper function $\varphi\colon \mathcal{X} \to \,]-\infty, +\infty]$ with conjugate φ^*, and define the associated *Fenchel–Young function* by

$$L_\varphi\colon \mathcal{X} \times \mathcal{X}^* \to \,]-\infty, +\infty]\colon (x, u^*) \mapsto \varphi(x) + \varphi^*(u^*) - \langle x, u^* \rangle. \tag{1.4}$$

As is well known [30, Section 10],

$$L_\varphi \geqslant 0 \quad \text{and} \quad \operatorname{gra} \partial\varphi = \{(x, u^*) \in \mathcal{X} \times \mathcal{X}^* \mid L_\varphi(x, u^*) = 0\}, \tag{1.5}$$

and one may likewise ask whether better lower bounds can be found. These two questions are not only of theoretical interest but they also impact concrete applications in areas such as inverse problems [1], evolution inclusions [3], optimal transportation [20, 21], and machine learning [11, 35].

Fix $(x, u^*) \in \mathcal{X} \times \mathcal{X}^*$. Then, in view of (1.1), given $(y, y^*) \in \operatorname{gra} A$, the inequality $H_A(x, u^*) \geqslant \langle x - y, y^* - u^* \rangle$ furnishes a trivial lower bound. To make it exploitable, (y, y^*) should depend on (x, u^*) in a suitable way. The first instance of such a lower bound we have found in the literature is the following.

Proposition 1.1 ([34, Lemma 2.3]). *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\Delta\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be the* duality mapping *of $\mathcal{X}$, i.e.,*

$$(\forall x \in \mathcal{X}) \quad \Delta(x) = \{x^* \in \mathcal{X}^* \mid \|x\|^2 = \langle x, x^* \rangle = \|x^*\|^2\}. \tag{1.6}$$

Then there exists $(z, z^*) \in \operatorname{gra} A$ such that $z^* - u^* \in \Delta(x - z)$ and $H_A(x, u^*) \geqslant \|x - z\|^2$.

If $\mathcal{X}$ is a Hilbert space (identified with its dual), then $\Delta = \mathrm{Id}$ and the inclusion $z^* - u^* \in \Delta(x - z)$ in Proposition 1.1 becomes $x + u^* - z = z^* \in Az$, i.e., $z = (\mathrm{Id} + A)^{-1}(x + u^*) = J_A(x + u^*)$ since the *resolvent* $J_A = (\mathrm{Id} + A)^{-1}$ is single-valued by [9, Corollary 23.11(i)]. We then deduce at once from Proposition 1.1 that

$$H_A(x, u^*) \geqslant \|x - J_A(x + u^*)\|^2. \tag{1.7}$$

This inequality appears explicitly in [3, Equation (1)], where it is also derived from [34, Lemma 2.3]. Unaware of these results, Carlier recently proposed in [20] a parametrized version of (1.7) along with a lower bound for (1.4), with elegant applications to convex analysis and transportation theory (see also [10] for further results in Hilbert spaces).

Proposition 1.2. *Suppose that $\mathcal{X}$ is a Hilbert space, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}$, and let $\gamma \in \,]0, +\infty[$. Then the following hold:*

(i) [20, Section 2] *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}}$ be maximally monotone. Then*

$$H_A(x, u^*) \geqslant \frac{\left\|x - J_{\gamma A}(x + \gamma u^*)\right\|^2}{\gamma}. \tag{1.8}$$

(ii) [20, Lemma 1.1] *Let $\varphi\colon \mathcal{X} \to \,]-\infty, +\infty]$ be a proper lower semicontinuous convex function and let* prox_φ *be its* proximity operator, *i.e.,* $\mathrm{prox}_\varphi = J_{\partial\varphi}$. *Then*

$$L_\varphi(x, u^*) \geqslant \frac{\left\|x - \mathrm{prox}_{\gamma\varphi}(x + \gamma u^*)\right\|^2}{\gamma}. \tag{1.9}$$

Following [20], an extension of Proposition 1.2(i) to reflexive real Banach spaces was proposed in [18] in the following form (this result is stated with the condition $\mathrm{dom}\, H_W = \mathcal{X} \times \mathcal{X}^*$ in [18], but the weaker, more checkable condition $\mathrm{dom}\, W = \mathcal{X}$ suffices; see Example 5.12).

Proposition 1.3 ([18, Theorem 1]). *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $W\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \,]0, +\infty[$. Suppose that W is α-strongly monotone for some $\alpha \in \,]0, +\infty[$, and* $\mathrm{dom}\, W = \mathcal{X}$. *Let $x^* \in Wx$ and set $z = (W + \gamma A)^{-1}(x^* + \gamma u^*)$. Then $H_A(x, u^*) \geqslant \alpha\|x - z\|^2/\gamma$.*

The main objective of the present paper is to derive new lower bounds on the Haraux and Fenchel–Young functions in reflexive real Banach spaces under minimal assumptions on the set-valued operator A in (1.1) and the function φ in (1.4), respectively. As seen above, the basic inequality (1.7) in Hilbert spaces can be deduced from Proposition 1.1. It can also be deduced from Proposition 1.3 with $W = \mathrm{Id}$. However, these propositions yield in general different conclusions. As we shall see, two notions of resolvent for set-valued operators implicitly underlie these inequalities: metric resolvents in Proposition 1.1, and warped resolvents in Proposition 1.3. In the Hilbertian setting, metric resolvents coincide with instances of warped resolvents and we recover Proposition 1.2.

In Section 2, we introduce our notation. Lower bounds on the Haraux function of general set-valued operators based on metric resolvents are derived in Section 3 and lower bounds on the Fenchel–Young function of proper functions based on metric proximity operators are derived in Section 4. In Sections 5 and 6, we provide alternative bounds based on warped resolvents for the Haraux function and on warped proximity operators for the Fenchel–Young function. These bounds are new even in the basic setting of maximally monotone operators and lower semicontinuous convex functions in Euclidean spaces, in which case we show that they can be more easily computable and sharper than those produced by Proposition 1.2. Section 7 proposes applications to composite monotone inclusions and Section 8 concludes the paper with some potential directions for future work.

§2. Notation

Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and let $A^{-1}\colon \mathcal{X}^* \to 2^{\mathcal{X}}\colon x^* \mapsto \{x \in \mathcal{X} \mid x^* \in Ax\}$ be its *inverse*. The *domain* of A is $\operatorname{dom} A = \{x \in \mathcal{X} \mid Ax \neq \varnothing\}$, the *range* of A is $\operatorname{ran} A = \bigcup_{x\in\operatorname{dom} A} Ax$, and the set of zeros of A is $\operatorname{zer} A = \{x \in \mathcal{X} \mid 0 \in Ax\}$. We say that A is *monotone* if

$$(\forall (x,x^*) \in \operatorname{gra} A)(\forall (y,y^*) \in \operatorname{gra} A) \quad \langle x-y, x^*-y^*\rangle \geqslant 0 \tag{2.1}$$

and *maximally monotone* if

$$(\forall (x,x^*) \in \mathcal{X}\times\mathcal{X}^*) \quad \big[\, (x,x^*) \in \operatorname{gra} A \;\Leftrightarrow\; (\forall (y,y^*) \in \operatorname{gra} A)\;\; \langle x-y, x^*-y^*\rangle \geqslant 0 \,\big]. \tag{2.2}$$

Let $\phi\colon [0,+\infty[\; \to [0,+\infty]$ be increasing and vanishing only at 0. Then A is *ϕ-uniformly monotone* if

$$(\forall (x,x^*) \in \operatorname{gra} A)(\forall (y,y^*) \in \operatorname{gra} A) \quad \langle x-y, x^*-y^*\rangle \geqslant \phi(\|x-y\|). \tag{2.3}$$

In particular, if $\phi = \alpha|\cdot|^2$ for some $\alpha \in \,]0,+\infty[$, then A is *α-strongly monotone*, that is,

$$(\forall (x,x^*) \in \operatorname{gra} A)(\forall (y,y^*) \in \operatorname{gra} A) \quad \langle x-y, x^*-y^*\rangle \geqslant \alpha\|x-y\|^2. \tag{2.4}$$

If A is monotone and $\operatorname{dom} A\times\operatorname{ran} A \subset \operatorname{dom} H_A$, then it is 3^* *monotone*. At last, A is *injective* if $(\forall x \in \mathcal{X})$ $(\forall y \in \mathcal{X})$ $Ax \cap Ay \neq \varnothing \Rightarrow x = y$.

Let $\varphi\colon \mathcal{X} \to [-\infty,+\infty]$. The *domain* of φ is $\operatorname{dom}\varphi = \{x \in \mathcal{X} \mid \varphi(x) < +\infty\}$ and the *conjugate* of φ is the function $\varphi^*\colon \mathcal{X}^* \to [-\infty,+\infty]\colon x^* \mapsto \sup_{x\in\mathcal{X}}(\langle x,x^*\rangle - \varphi(x))$. Further, φ is *cofinite* if $\operatorname{dom}\varphi^* = \mathcal{X}^*$ and *supercoercive* if $\lim_{\|x\|\to+\infty} \varphi(x)/\|x\| = +\infty$. Now suppose that $\varphi\colon \mathcal{X} \to \,]-\infty,+\infty]$ is *proper*, i.e., $\operatorname{dom}\varphi \neq \varnothing$. The *subdifferential* of φ is the operator

$$\partial\varphi\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon x \mapsto \big\{x^* \in \mathcal{X}^* \mid (\forall y \in \mathcal{X})\; \langle y-x, x^*\rangle + \varphi(x) \leqslant \varphi(y)\big\}. \tag{2.5}$$

We denote by $\Gamma_0(\mathcal{X})$ the class of proper lower semicontinuous convex functions from $\mathcal{X}$ to $]-\infty,+\infty]$. Let $f \in \Gamma_0(\mathcal{X})$. If f is Gateaux differentiable on $\operatorname{int}\operatorname{dom} f \neq \varnothing$, then the *Bregman distance* associated with f is

$$\begin{aligned} D_f\colon \mathcal{X}\times\mathcal{X} &\to [0,+\infty] \\ (x,y) &\mapsto \begin{cases} f(x) - f(y) - \langle x-y, \nabla f(y)\rangle, & \text{if } \; y \in \operatorname{int}\operatorname{dom} f;\\ +\infty, & \text{otherwise.}\end{cases}\end{aligned} \tag{2.6}$$

Finally, f is a *Legendre function* if it is essentially smooth in the sense that ∂f is both locally bounded and single-valued on its domain, and essentially strictly convex in the sense that ∂f^* is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom}\partial f$ [7].

§3. Lower bounds on the Haraux function based on metric resolvents

The Voisei–Zălinescu inequality [39, Theorem 2.6] gives a lower bound on the Haraux function of a maximally monotone operator in terms of the distance to its graph. The main result of this section is a sharpening of this inequality. As in Proposition 1.1, the duality mapping of (1.6) plays a central role via the following notion of a metric resolvent (see [33, Definition 3.4] and the references therein).

Definition 3.1. Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and let Δ be the duality mapping of $\mathcal{X}$. Then the *metric resolvent* of A is

$$R_A\colon \mathcal{X} \to 2^{\mathcal{X}}\colon x \mapsto \big\{z \in \mathcal{X} \mid 0 \in Az + \Delta(z - x)\big\}. \tag{3.1}$$

Let us recall a few facts about duality mappings.

Lemma 3.2. *Let Δ be the duality mapping of $\mathcal{X}$. Then the following hold:*

(i) [23, Theorem I.4.4] $\Delta = \partial\|\cdot\|^2/2$.
(ii) [23, Corollary V.2.6] *Δ is maximally monotone.*
(iii) [23, Proposition I.4.7(c)] *Let $x \in \mathcal{X}$. Then $\Delta(-x) = -\Delta(x)$.*
(iv) [23, Theorem II.1.8] *Suppose that $\mathcal{X}$ is strictly convex. Then Δ is strictly monotone.*

Example 3.3. Let C be a nonempty closed convex subset of $\mathcal{X}$, let proj_C be the *metric projection* operator onto C, i.e.,

$$\operatorname{proj}_C\colon \mathcal{X} \to 2^C\colon x \mapsto \big\{z \in C \mid (\forall y \in C)\ \ \|x - z\| \leqslant \|x - y\|\big\}, \tag{3.2}$$

and let N_C be the normal cone operator of C, i.e.,

$$N_C\colon \mathcal{X} \to 2^{\mathcal{X}^*}\colon x \mapsto \begin{cases} \big\{x^* \in \mathcal{X}^* \;\big|\; (\forall y \in C)\ \langle y - x, x^*\rangle \leqslant 0\big\}, & \text{if } x \in C;\\ \varnothing, & \text{otherwise.}\end{cases} \tag{3.3}$$

Then it follows from [28, Remarque 8.1.5a], Lemma 3.2(i), and Definition 3.1 that $\operatorname{proj}_C = R_{N_C}$.

Notation 3.4. Let $\gamma \in \left]0, +\infty\right[$. We define a norm by

$$|||\cdot|||_\gamma\colon \mathcal{X} \times \mathcal{X}^* \to \left[0, +\infty\right[\colon (x, x^*) \mapsto \sqrt{\|x\|^2/\gamma + \gamma\|x^*\|^2} \tag{3.4}$$

and denote the associated distance function to a set $\boldsymbol{C} \subset \mathcal{X} \times \mathcal{X}^*$ by

$$d_{\boldsymbol{C},\gamma}\colon \mathcal{X} \times \mathcal{X}^* \to [0, +\infty]\colon (x, x^*) \mapsto \inf_{(y,y^*)\in\boldsymbol{C}} |||(x, x^*) - (y, y^*)|||_\gamma. \tag{3.5}$$

The proposed lower bounds are as follows.

Proposition 3.5. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Suppose that $z \in R_{\gamma(A-u^*)}x$. Then, using Notation 3.4,*

$$H_A(x, u^*) \geqslant \frac{\|x - z\|^2}{\gamma} \geqslant \frac{1}{2} d^2_{\operatorname{gra} A,\gamma}(x, u^*). \tag{3.6}$$

Proof. Since $z \in R_{\gamma(A-u^*)}x$, it follows from (3.1) that there exists $z^* \in Az$ such that $0 \in \gamma(z^* - u^*) + \Delta(z - x)$. Consequently, $(z, z^*) \in \operatorname{gra} A$ and, appealing to Lemma 3.2(iii), we obtain

$$\gamma(z^* - u^*) \in \Delta(x - z). \tag{3.7}$$

In turn, (1.1) and (1.6) yield

$$H_A(x, u^*) \geqslant \langle x - z, z^* - u^*\rangle = \frac{\langle x - z, \gamma(z^* - u^*)\rangle}{\gamma} = \frac{\|x - z\|^2}{\gamma}. \tag{3.8}$$

Next, we derive from (3.7), (1.6), and (3.5) that

$$\begin{aligned}\frac{\|x-z\|^2}{\gamma} &= \frac{\|x-z\|^2+\gamma^2\|u^*-z^*\|^2}{2\gamma}\\ &= \frac{1}{2}|||(x,u^*)-(z,z^*)|||_\gamma^2\\ &\geqslant \frac{1}{2}\inf_{(y,y^*)\in\operatorname{gra}A}|||(x,u^*)-(y,y^*)|||_\gamma^2\\ &= \frac{1}{2}d^2_{\operatorname{gra}A,\gamma}(x,u^*),\end{aligned} \tag{3.9}$$

which yields the rightmost inequality. ☐

Theorem 3.6. *Let $A\colon \mathcal{X}\to 2^{\mathcal{X}^*}$ be maximally monotone, let $x\in\mathcal{X}$, let $u^*\in\mathcal{X}^*$, and let $\gamma\in\,]0,+\infty[$. Then, using Notation 3.4, the following hold:*

(i) $R_{\gamma(A-u^*)}x\neq\varnothing$.

(ii) *Let $z\in R_{\gamma(A-u^*)}x$. Then*

$$H_A(x,u^*)\geqslant\frac{\|x-z\|^2}{\gamma}\geqslant\frac{1}{2}d^2_{\operatorname{gra}A,\gamma}(x,u^*). \tag{3.10}$$

(iii) *Suppose that $\mathcal{X}$ is strictly convex. Then $R_{\gamma(A-u^*)}$ is single-valued and*

$$H_A(x,u^*)\geqslant\frac{\|x-R_{\gamma(A-u^*)}x\|^2}{\gamma}\geqslant\frac{1}{2}d^2_{\operatorname{gra}A,\gamma}(x,u^*). \tag{3.11}$$

Proof. (i): By [37, Theorem 10.6], $(x,\gamma u^*)\in\mathcal{X}\times\mathcal{X}^*=\operatorname{gra}(\gamma A)+\operatorname{gra}(-\Delta)$. Therefore, there exists $(z,z^*)\in\operatorname{gra}A$ such that $(x-z,\gamma u^*-\gamma z^*)\in\operatorname{gra}(-\Delta)$, hence $\gamma(z^*-u^*)\in\Delta(x-z)$, which implies that $0\in\gamma(z^*-u^*)+\Delta(z-x)\subset\gamma(A-u^*)z+\Delta(z-x)$. In view of Definition 3.1, $z\in R_{\gamma(A-u^*)}x$.

(ii): This follows from (i) and Proposition 3.5.

(iii): Taking (ii) into account, it is enough to show that $R_{\gamma(A-u^*)}x$ contains at most one point. Suppose that $\{z_1,z_2\}\subset R_{\gamma(A-u^*)}x$. Then we infer from Definition 3.1 that there exist $w_1^*\in\Delta(z_1-x)$ and $w_2^*\in\Delta(z_2-x)$ such that $(z_1,-w_1^*)\in\operatorname{gra}A$ and $(z_2,-w_2^*)\in\operatorname{gra}A$. By monotonicity of A,

$$\langle(z_1-x)-(z_2-x),w_1^*-w_2^*\rangle=\langle z_1-z_2,w_1^*-w_2^*\rangle\leqslant 0. \tag{3.12}$$

However, Lemma 3.2(ii) forces

$$\langle(z_1-x)-(z_2-x),w_1^*-w_2^*\rangle\geqslant 0. \tag{3.13}$$

Thus, $\langle(z_1-x)-(z_2-x),w_1^*-w_2^*\rangle=0$ and, since Lemma 3.2(iv) asserts that Δ is strictly monotone, we conclude that $z_1-x=z_2-x$, i.e., that $z_1=z_2$. ☐

Remark 3.7. Theorem 3.6 extends existing results as follows:

(i) For $\gamma=1$, it follows in particular from Theorem 3.6(ii) that, if $z\in R_{A-u^*}x$, then $H_A(x,u^*)\geqslant\|x-z\|^2$. This conclusion is identical to that of [34, Lemma 2.3] (see Proposition 1.1). Indeed, arguing as in the proof of Theorem 3.6(i), the inclusion $z\in R_{A-u^*}x$ secures the existence of $(z,z^*)\in\operatorname{gra}A$ such that $z^*-u^*\in\Delta(x-z)$.

(ii) Regarding Theorem 3.6(ii), the smaller lower bound

$$H_A(x,u^*) \geqslant \frac{1}{4}d^2_{\operatorname{gra} A,1}(x,u^*) \tag{3.14}$$

was established in [39, Theorem 2.6] using more technical arguments. Our simple proof improves it by a factor 2, and for any $\gamma \in \left]0,+\infty\right[$.

(iii) Regarding Theorem 3.6(ii), in the special case when $\mathcal{X}$ is Hilbertian and $\gamma = 1$, the inequality

$$H_A(x,u^*) \geqslant \frac{1}{2}d^2_{\operatorname{gra} A,1}(x,u^*) \tag{3.15}$$

appears in [18, Theorem 4].

(iv) As pointed out in [12, Exercise 9.7.11], the constant $1/2$ in Theorem 3.6(ii) is the best possible to the extent that, if $\mathcal{X}$ is Hilbertian, $A = \mathrm{Id}$, and $\gamma = 1$, then the inequalities become equalities.

(v) Suppose that $\mathcal{X}$ is Hilbertian. Then $\Delta = \mathrm{Id}$ and it follows from Definition 3.1 that $R_{\gamma A}$ coincides with the usual resolvent $J_{\gamma A} = (\mathrm{Id} + \gamma A)^{-1}$ of [9, Definition 23.1]. In this context, the inequality $H_A(x,u^*) \geqslant \|x - J_{\gamma(A-u^*)}x\|^2/\gamma$ provided by Theorem 3.6(iii) is easily seen to be equivalent to the inequality $H_A(x,u^*) \geqslant \left\|x - J_{\gamma A}(x+\gamma u^*)\right\|^2/\gamma$, which is precisely that established in [20, Section 2] (see Proposition 1.2(i)).

§4. Lower bounds on the Fenchel–Young function based on metric proximity operators

We turn our attention to the case when A is a subdifferential operator to obtain lower bounds on the Fenchel–Young function of (1.4). The strategy is to exploit the conclusions of Section 3 via the following inequality which, in the case of Hilbert spaces, appears in [13, p. 167] (if φ is convex, it is also a consequence of [26, Theorem 3.7] and (1.2)). For completeness, we record a proof in our context.

Lemma 4.1. *Let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be proper. Then $L_\varphi \geqslant H_{\partial\varphi}$.*

Proof. Let $(x,u^*) \in \mathcal{X}\times\mathcal{X}^*$ and $(y,y^*) \in \operatorname{gra}\partial\varphi$. Then, since $y \in \operatorname{dom}\varphi$, (2.5) yields

$$\begin{aligned}\varphi(x)+\varphi^*(u^*)-\langle x,u^*\rangle &\geqslant \varphi(x)+\langle y,u^*\rangle-\varphi(y)-\langle x,u^*\rangle\\ &\geqslant \langle x-y,y^*\rangle+\varphi(y)+\langle y,u^*\rangle-\varphi(y)-\langle x,u^*\rangle\\ &= \langle x-y,y^*-u^*\rangle.\end{aligned} \tag{4.1}$$

In view of (1.1) and (1.4), this shows that $L_\varphi(x,u^*) \geqslant H_{\partial\varphi}(x,u^*)$. ☐

Combining Proposition 3.5 and Lemma 4.1 yields the following set of inequalities in terms of the metric resolvent of Definition 3.1.

Proposition 4.2. *Let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be proper, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0,+\infty\right[$. Suppose that $z \in R_{\partial(\gamma(\varphi-u^*))}x$. Then, using Notation 3.4,*

$$L_\varphi(x,u^*) \geqslant H_{\partial\varphi}(x,u^*) \geqslant \frac{\|x-z\|^2}{\gamma} \geqslant \frac{1}{2}d^2_{\operatorname{gra}\partial\varphi,\gamma}(x,u^*). \tag{4.2}$$

To refine these inequalities, let us recall the notion of a metric proximity operator, first introduced by Moreau [29] in Hilbert spaces.

Definition 4.3. Let $\varphi \in \Gamma_0(\mathcal{X})$. Then the *metric proximity operator* of φ is

$$\operatorname{prox}_\varphi \colon \mathcal{X} \to 2^{\mathcal{X}} \colon x \mapsto \underset{y \in \mathcal{X}}{\operatorname{Argmin}} \left(\varphi(y) + \frac{1}{2}\|x - y\|^2 \right). \tag{4.3}$$

The following theorem establishes lower bounds on L_φ for $\varphi \in \Gamma_0(\mathcal{X})$.

Theorem 4.4. *Let $\varphi \in \Gamma_0(\mathcal{X})$, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Then, using Notation 3.4, the following hold:*

(i) $\operatorname{prox}_{\gamma(\varphi - u^*)} x \neq \varnothing$.

(ii) *Let $z \in \operatorname{prox}_{\gamma(\varphi - u^*)} x$. Then*

$$L_\varphi(x, u^*) \geqslant H_{\partial\varphi}(x, u^*) \geqslant \frac{\|x - z\|^2}{\gamma} \geqslant \frac{1}{2} d^2_{\operatorname{gra}\partial\varphi, \gamma}(x, u^*). \tag{4.4}$$

(iii) *Suppose that $\mathcal{X}$ is strictly convex. Then $\operatorname{prox}_{\gamma(\varphi - u^*)}$ is single-valued and*

$$L_\varphi(x, u^*) \geqslant H_{\partial\varphi}(x, u^*) \geqslant \frac{\|x - \operatorname{prox}_{\gamma(\varphi - u^*)} x\|^2}{\gamma} \geqslant \frac{1}{2} d^2_{\operatorname{gra}\partial\varphi, \gamma}(x, u^*). \tag{4.5}$$

Proof. First, [40, Proposition 47.F(1)] asserts that $\partial\varphi$ is maximally monotone. Next, we derive from Definition 4.3, Fermat's rule [40, Proposition 47.12], the subdifferential sum rule [40, Theorem 47.B], Lemma 3.2(i), and Definition 3.1 that

$$\begin{aligned}
\operatorname{prox}_{\gamma(\varphi - u^*)} x &= \underset{y \in \mathcal{X}}{\operatorname{Argmin}} \left(\gamma\big(\varphi(y) - \langle y, u^* \rangle\big) + \frac{1}{2}\|x - y\|^2 \right) \\
&= \operatorname{zer} \partial\Big(\gamma\big(\varphi - u^*\big) + \frac{1}{2}\|\cdot - x\|^2 \Big) \\
&= \operatorname{zer}\big(\gamma\big(\partial\varphi - u^*\big) + \Delta(\cdot - x) \big) \\
&= \big\{ z \in \mathcal{X} \;\big|\; 0 \in \gamma(\partial\varphi - u^*)z + \Delta(z - x) \big\} \\
&= R_{\gamma(\partial\varphi - u^*)} x.
\end{aligned} \tag{4.6}$$

The claims are therefore consequences of Theorem 3.6, where $A = \partial\varphi$, and Lemma 4.1. ☐

Remark 4.5. Suppose that $\mathcal{X}$ is Hilbertian. Then, as in Remark 3.7(v), the inequality $L_\varphi(x, u^*) \geqslant \|x - \operatorname{prox}_{\gamma(\varphi - u^*)} x\|^2/\gamma$ from Theorem 4.4(iii) is equivalent to $L_\varphi(x, u^*) \geqslant \|x - \operatorname{prox}_{\gamma\varphi}(x + \gamma u^*)\|^2/\gamma$, which is precisely the lower bound established in [20, Lemma 1.1] (see Proposition 1.2(ii)).

§5. Lower bounds on the Haraux function based on warped resolvents

We derive lower bounds on the Haraux function of a general set-valued operator $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ in terms of an auxiliary set-valued operator $W \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ through the notion of a warped resolvent. The results are then specialized to the case when W is at most single-valued and, in particular, when it is the gradient of a Legendre function, which gives rise to lower bounds in terms of Bregman distances. Several examples illustrate these new bounds and an application to the asymptotic behavior of families of set-valued operators is provided. Our analysis relies on the following notion of a warped resolvent, which was introduced in the case of at-most single-valued kernels in [15, Definition 1.1].

Definition 5.1. Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $K\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Then the *warped resolvent* of A with kernel K is $J_A^K = (K+A)^{-1} \circ K$.

Proposition 5.2. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $W\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $x \in \mathcal{X}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0,+\infty\right[$. Then*

$$J^W_{\gamma(A-u^*)}x = (W+\gamma A)^{-1}(Wx+\gamma u^*). \tag{5.1}$$

Proof. Let $z \in \mathcal{X}$. Then

$$\begin{aligned} z \in J^W_{\gamma(A-u^*)}x \quad &\Leftrightarrow \quad z \in \big(W+\gamma(A-u^*)\big)^{-1}(Wx) \\ &\Leftrightarrow \quad (\exists\, x^* \in Wx) \quad z \in \big(W+\gamma(A-u^*)\big)^{-1}x^* \\ &\Leftrightarrow \quad (\exists\, x^* \in Wx) \quad x^*+\gamma u^* \in Wz+\gamma Az \\ &\Leftrightarrow \quad (\exists\, x^* \in Wx) \quad z \in (W+\gamma A)^{-1}(x^*+\gamma u^*) \\ &\Leftrightarrow \quad z \in (W+\gamma A)^{-1}(Wx+\gamma u^*), \end{aligned} \tag{5.2}$$

which proves (5.1). ☐

Proposition 5.3. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $W\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0,+\infty\right[$. Suppose that $x \in \operatorname{dom} W$ and that $z \in J^W_{\gamma(A-u^*)}x$. Then the following hold:*

(i) *There exist $x^* \in Wx$ and $z^* \in Wz$ such that $H_A(x,u^*) \geqslant \langle x-z, x^*-z^*\rangle/\gamma$.*

(ii) *Suppose that W is ϕ-uniformly monotone. Then $H_A(x,u^*) \geqslant \phi(\|x-z\|)/\gamma$.*

(iii) *Suppose that W is α-strongly monotone. Then $H_A(x,u^*) \geqslant \alpha\|x-z\|^2/\gamma$.*

(iv) *Suppose that W is at most single-valued. Then $H_A(x,u^*) \geqslant \langle x-z, Wx-Wz\rangle/\gamma$.*

(v) *Let $f \in \Gamma_0(\mathcal{X})$ be Gateaux differentiable on* $\operatorname{dom}\nabla f = \operatorname{int}\operatorname{dom} f$ *and suppose that $W=\nabla f$. Then*

$$H_A(x,u^*) \geqslant \frac{D_f(x,z)+D_f(z,x)}{\gamma}. \tag{5.3}$$

Proof. (i): Since Proposition 5.2 asserts that $z \in J^W_{\gamma(A-u^*)}x = (W+\gamma A)^{-1}(Wx+\gamma u^*)$, there exists $x^* \in Wx$ such that $x^*+\gamma u^* \in Wz+\gamma Az$, from which we deduce that $z \in \operatorname{dom} W \cap \operatorname{dom} A$ and that there exists $z^* \in Wz$ such that $x^*+\gamma u^* \in z^*+\gamma Az$. In turn,

$$\left(z, \frac{x^*-z^*}{\gamma}+u^*\right) \in \operatorname{gra} A. \tag{5.4}$$

Consequently,

$$H_A(x,u^*) = \sup_{(y,y^*)\in\operatorname{gra} A} \langle x-y, y^*-u^*\rangle \geqslant \left\langle x-z, \frac{x^*-z^*}{\gamma}+u^*-u^*\right\rangle = \frac{\langle x-z, x^*-z^*\rangle}{\gamma}. \tag{5.5}$$

(ii): This follows from (i) and (2.3).
(iii): Take $\phi = \alpha|\cdot|^2$ in (ii).
(iv): An immediate consequence of (i).
(v): This follows from (iv) and (2.6). ☐

Let us characterize the situation in which the point z in Proposition 5.3(iv) is x itself .

Proposition 5.4. *In the setting of Proposition 5.3(iv), consider the following statements:*

[a] $x \in J^W_{\gamma(A-u^*)}x$.

[b] $(x,u^*) \in \operatorname{gra} A$.

[c] $H_A(x,u^*) = 0$.

Then the following hold:

(i) [a]⇔[b].

(ii) *Suppose that A is monotone. Then* [b]⇒[c].

(iii) *Suppose that A is maximally monotone. Then* [b]⇔[c].

Proof. (i): We have $x \in J^W_{\gamma(A-u^*)}x \Leftrightarrow x \in (W+\gamma(A-u^*))^{-1}(Wx) \Leftrightarrow Wx \in Wx+\gamma(Ax-u^*) \Leftrightarrow u^* \in Ax$.

(ii): By monotonicity, $(\forall (y,y^*) \in \operatorname{gra} A)$ $\langle x-y, y^*-u^*\rangle \leqslant 0$. Therefore, $H_A(x,u^*) \leqslant 0$. At the same time, $H_A(x,u^*) = \sup_{(y,y^*)\in \operatorname{gra} A}\langle x-y, y^*-u^*\rangle \geqslant \langle x-x, u^*-u^*\rangle = 0$.

(iii): See (1.3). ☐

The following result ensures the existence of the point z in Proposition 5.3.

Proposition 5.5. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $W\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0,+\infty\right[$. Suppose that $x \in \operatorname{dom} W$ and that one of the following holds:*

(i) *There exists $x^* \in Wx$ such that $x^* + \gamma u^* \in \operatorname{ran}(W+\gamma A)$.*

(ii) *$W+\gamma A$ is surjective.*

(iii) *$W+\gamma A$ is maximally monotone and $(W+\gamma A)^{-1}$ is locally bounded at every point in $\mathcal{X}^*$.*

(iv) *$W+\gamma A$ is maximally monotone and one of the following is satisfied:*

  (a) *$\operatorname{dom} W \cap \operatorname{dom} A$ is bounded.*

  (b) *$\operatorname{dom} W \cap \operatorname{dom} A$ is unbounded and*

$$\lim_{\substack{y\in \operatorname{dom} W\cap \operatorname{dom} A\\ \|y\|\to+\infty}} \left(\inf_{y^*\in Wy+\gamma Ay} \|y^*\|\right) = +\infty. \tag{5.6}$$

(v) *$W+\gamma A$ is maximally monotone and ϕ-uniformly monotone with $\phi(t)/t \to +\infty$ as $t \to +\infty$.*

(vi) *$W+\gamma A$ is maximally monotone and α-strongly monotone for some $\alpha \in \left]0,+\infty\right[$.*

(vii) *W and A are monotone, $W+\gamma A$ is maximally monotone, $\operatorname{ran} W + \gamma \operatorname{ran} A = \mathcal{X}^*$, and one of the following is satisfied:*

  (a) *W and A are 3^* monotone.*

  (b) *W is 3^* monotone and $\operatorname{dom} A \subset \operatorname{dom} W$.*

  (c) *A is 3^* monotone and $\operatorname{dom} W \subset \operatorname{dom} A$.*

(viii) *$W = \partial f$, with $f \in \Gamma_0(\mathcal{X})$, A is monotone, $\partial f + \gamma A$ is maximally monotone, $\operatorname{dom} \partial f^* + \gamma \operatorname{ran} A = \mathcal{X}^*$, and either $\operatorname{dom} A \subset \operatorname{dom} \partial f$ or A is 3^* monotone.*

(ix) *W and A are monotone, $W+\gamma A$ is maximally monotone, and one of the following is satisfied:*

  (a) *W is 3^* monotone and surjective.*

  (b) *A is 3^* monotone and surjective.*

Then $J^W_{\gamma(A-u^*)}x \neq \varnothing$.

Proof. For convenience, we set $M = W + \gamma A$.

(i): Since $x^* + \gamma u^* \in \operatorname{ran} M = \operatorname{dom} M^{-1}$, we have $M^{-1}(x^* + \gamma u^*) \neq \varnothing$. Therefore, by Proposition 5.2, $J^W_{\gamma(A-u^*)} x = M^{-1}(Wx + \gamma u^*) \neq \varnothing$.

(ii)⇒(i): Indeed, $\operatorname{ran} M = \mathcal{X}^*$.

(iii)⇒(ii): This follows from the Brézis–Browder surjectivity theorem [41, Theorem 32.G].

(iv)⇒(ii): See [41, Corollary 32.35].

(v)⇒(iv): In view of (iv), we assume that $\operatorname{dom} M$ is unbounded and show that

$$\lim_{\substack{y \in \operatorname{dom} M \\ \|y\| \to +\infty}} \left(\inf_{y^* \in My} \|y^*\| \right) = +\infty. \tag{5.7}$$

For this purpose, fix $(v, v^*) \in \operatorname{gra} M$. Then (2.3) yields

$$(\forall (y, y^*) \in \operatorname{gra} M) \quad \|y - v\| \, \|y^* - v^*\| \geqslant \langle y - v, y^* - v^* \rangle \geqslant \phi(\|y - v\|). \tag{5.8}$$

Therefore

$$(\forall (y, y^*) \in \operatorname{gra} M) \quad y \neq v \;\Rightarrow\; \|y^*\| \geqslant \|y^* - v^*\| - \|v^*\| \geqslant \frac{\phi(\|y - v\|)}{\|y - v\|} - \|v^*\|. \tag{5.9}$$

Since the right-hand side goes to $+\infty$ as $\|y\| \to +\infty$, we obtain (5.7).

(vi)⇒(v): Take $\phi = \alpha |\cdot|^2$.

(vii)⇒(ii): It follows from the assumptions and the reflexive Banach space version [36, Theorem 2.2] of the Brézis–Haraux theorem [13, Théorèmes 3 et 4] that $\operatorname{int} \operatorname{ran} M = \operatorname{int}(\operatorname{ran} W + \gamma \operatorname{ran} A)$, which implies that $\operatorname{ran} M = \mathcal{X}^*$.

(viii)⇒(vii): Indeed, W is 3^* monotone by [41, Proposition 32.42] and $\operatorname{ran} W = \operatorname{ran} \partial f = \operatorname{dom}(\partial f)^{-1} = \operatorname{dom} \partial f^*$ by [40, Theorem 51.A(ii)].

(ix)⇒(ii): See [14, Corollary 11 and Remark 9(iv)]. ☐

Example 5.6. Let $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $W \colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, let $x \in \operatorname{dom} W$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Suppose that the cone generated by $\operatorname{dom} W - \operatorname{dom} A$ is a closed vector subspace of $\mathcal{X}$ and that W is ϕ-uniformly monotone with $\phi(t)/t \to +\infty$ as $t \to +\infty$. Let $z \in J^W_{\gamma(A-u^*)} x$. Then $H_A(x, u^*) \geqslant \phi(\|x - z\|)/\gamma$.

Proof. It follows from the Attouch–Riahi–Théra theorem [38, Corollary 32.3] that $W + \gamma A$ is maximally monotone. Additionally, $W + \gamma A$ is ϕ-uniformly monotone and therefore Proposition 5.5(v) guarantees that $J^W_{\gamma(A-u^*)} x \neq \varnothing$. The conclusion therefore follows from Proposition 5.3(ii). ☐

An important consequence of Propositions 5.3(iv) and 5.5 is the following.

Theorem 5.7. *Let $A \colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W \colon \mathcal{D} \to \mathcal{X}^*$, let $x \in \mathcal{D}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Suppose that $\mathcal{D} \cap \operatorname{dom} A \neq \varnothing$ and that one of properties (i)–(ix) in Proposition 5.5 is satisfied, together with one of the following:*

(i) *$W + \gamma A$ is injective.*

(ii) *$W + \gamma A$ is strictly monotone.*

(iii) *W is uniformly monotone and A is monotone.*

Set $z = J^W_{\gamma(A-u^*)} x$. Then $H_A(x, u^*) \geqslant \langle x - z, Wx - Wz \rangle / \gamma$.

Proof. Set $M = W + \gamma A$. In view of Propositions 5.3(iv), 5.5, and 5.2, it remains to show that the set $M^{-1}(Wx + \gamma u^*) = J^W_{\gamma(A-u^*)}x$ is a singleton.

(i): Let (x^*, x_1) and (x^*, x_2) be points in $\operatorname{gra} M^{-1}$. Then $x^* \in Mx_1 \cap Mx_2$ and therefore, by injectivity, $x_1 = x_2$. Thus, M^{-1} is at most single-valued.

(ii)⇒(i): Let $(x_1, x_2) \in \mathcal{X} \times \mathcal{X}$ be such that there exists $x^* \in Mx_1 \cap Mx_2$. Then (x_1, x^*) and (x_2, x^*) lie in $\operatorname{gra} M$ and $\langle x_1 - x_2, x^* - x^* \rangle = 0$. The strict monotonicity of M then forces $x_1 = x_2$.

(iii)⇒(ii): Since W is strictly monotone and γA is monotone, $W + \gamma A$ is strictly monotone. □

Example 5.8. Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$, let $x \in \mathcal{D}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Suppose that A and W are maximally monotone, that the cone generated by $\mathcal{D} - \operatorname{dom} A$ is a closed vector subspace of $\mathcal{X}$, and that W is ϕ-uniformly monotone with $\phi(t)/t \to +\infty$ as $t \to +\infty$. Set $z = J^W_{\gamma(A-u^*)}x$. Then $H_A(x, u^*) \geqslant \phi(\|x - z\|)/\gamma$.

Proof. The Attouch–Riahi–Théra theorem [38, Corollary 32.3] guarantees that $W + \gamma A$ is maximally monotone. Since it is also ϕ-uniformly monotone, condition (v) of Proposition 5.5 is satisfied. The conclusion therefore follows from Theorem 5.7(iii) and (2.3). □

Our next result establishes a lower bound in terms of Bregman distances. Here, W is the gradient of a Legendre function $f \in \Gamma_0(\mathcal{X})$ and the warped resolvent $J^{\nabla f}_A$ becomes the Bregman resolvent of A studied in [8, Section 3.3].

Proposition 5.9. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone, let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $x \in \operatorname{int} \operatorname{dom} f$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0, +\infty\right[$. Suppose that $(\operatorname{int} \operatorname{dom} f) \cap \operatorname{dom} A \neq \varnothing$ and that one of the following holds:*

(i) $\gamma u^* + \operatorname{int} \operatorname{dom} f^* \subset \operatorname{ran}(\nabla f + \gamma A)$.

(ii) $(\operatorname{int} \operatorname{dom} f^*) + \gamma \operatorname{ran} A = \mathcal{X}^*$ *and one of the following is satisfied:*

   (a) *A is 3^* monotone.*

   (b) $\operatorname{dom} A \subset \operatorname{int} \operatorname{dom} f$.

Set $z = J^{\nabla f}_{\gamma(A-u^*)}x$. Then

$$H_A(x, u^*) \geqslant \frac{D_f(x, z) + D_f(z, x)}{\gamma}. \tag{5.10}$$

Proof. We apply Theorem 5.7 with $W = \nabla f$ and $\mathcal{D} = \operatorname{dom} W$. First, since f is essentially smooth, [7, Theorem 5.6] asserts that $\mathcal{D} = \operatorname{int} \operatorname{dom} f$. In addition, since W is maximally monotone and $(\operatorname{int} \operatorname{dom} W) \cap \operatorname{dom} A = (\operatorname{int} \operatorname{dom} f) \cap \operatorname{dom} A \neq \varnothing$, the Rockafellar sum theorem [41, Theorem 32.I] asserts that $W + \gamma A$ is maximally monotone. Moreover, since f is essentially strictly convex, it is strictly convex on the convex set $\mathcal{D} = \operatorname{int} \operatorname{dom} f$, which makes W strictly monotone [41, Proposition 25.10]. In turn, $W + \gamma A$ is strictly monotone. This shows that property (ii) in Theorem 5.7 is satisfied. On the other hand, since [7, Theorem 5.10] asserts that $\operatorname{ran} \nabla f = \operatorname{int} \operatorname{dom} f^*$, property (i) above implies property (i) in Proposition 5.5, while property (ii) above implies property (vii) in Proposition 5.5 since [7, Theorem 5.9(ii)] asserts that $\operatorname{dom} \partial f^* = \operatorname{dom} \nabla f^* = \operatorname{int} \operatorname{dom} f^*$, while [41, Proposition 32.42] asserts that W is 3^* monotone. Altogether, the conclusion follows from Theorem 5.7 and (2.6). □

Remark 5.10. A sufficient condition for the property $(\operatorname{int} \operatorname{dom} f^*) + \gamma \operatorname{ran} A = \mathcal{X}^*$ in Proposition 5.9(ii) to hold is that f be supercoercive, as this implies that f^* is cofinite [7, Theorem 3.4].

We now recover the existing results of Propositions 1.2(i) and 1.3.

Example 5.11. Suppose that $\mathcal{X}$ is a Hilbert space. Then we retrieve Proposition 1.2(i) by applying Proposition 5.9(ii)(b) with the Legendre function $f = \|\cdot\|^2/2$, which satisfies $f^* = f$, $\operatorname{int}\operatorname{dom} f = \operatorname{int}\operatorname{dom} f^* = \mathcal{X}$, and $D_f\colon \mathcal{X}\times\mathcal{X} \to \mathbb{R}\colon (u,v) \mapsto \|u-v\|^2/2$. Alternatively, we can apply Example 5.8 with $W = \mathrm{Id}$, which satisfies $\operatorname{dom} W = \mathcal{X}$ and is ϕ-uniformly monotone with $\phi = |\cdot|^2$.

Example 5.12. Using Proposition 5.2, we retrieve Proposition 1.3 as a special case of Example 5.6, where $\operatorname{dom} W = \mathcal{X}$ and $\phi = \alpha|\cdot|^2$.

The next example is an illustration of Proposition 5.9 for a maximally monotone operator which is not a subdifferential.

Example 5.13. Let $\mathcal{X} = \mathbb{R}^2$ be the standard Euclidean plane, let $\beta \in \left]0,+\infty\right[$, and let $\psi\colon \mathbb{R} \to \mathbb{R}$ be a Legendre function with a β-Lipschitzian derivative. Set

$$A\colon \mathcal{X} \to \mathcal{X}\colon (\xi_1,\xi_2) \mapsto \big(\beta\xi_1 - \psi'(\xi_1) - \xi_2, \xi_1 + \beta\xi_2 - \psi'(\xi_2)\big) \tag{5.11}$$

and consider the Legendre function

$$f\colon \mathcal{X} \to \mathbb{R}\colon (\xi_1,\xi_2) \mapsto \psi(\xi_1) + \psi(\xi_2). \tag{5.12}$$

As observed in [25, Remark 3.4], A is maximally monotone but it is not the subdifferential of a convex function. Now let $x = (\xi_1,\xi_2) \in \mathcal{X}$ and $u^* = (\mu_1^*,\mu_2^*) \in \mathcal{X}$. Since $\operatorname{ran}(\nabla f + A) = \mathcal{X}$, we derive from Propositions 5.9 and 5.2 that

$$\begin{aligned} z &= (\zeta_1,\zeta_2)\\ &= J^{\nabla f}_{A-u^*}x\\ &= (\nabla f + A)^{-1}\big(\nabla f(x) + u^*\big)\\ &= \left(\frac{\beta\big(\psi'(\xi_1)+\mu_1^*\big) + \psi'(\xi_2) + \mu_2^*}{1+\beta^2}, \frac{\beta\big(\psi'(\xi_2)+\mu_2^*\big) - \psi'(\xi_1) - \mu_1^*}{1+\beta^2}\right) \end{aligned} \tag{5.13}$$

is well defined and that

$$D_f(x,z) + D_f(z,x) = (\xi_1-\zeta_1)\big(\psi'(\xi_1)-\psi'(\zeta_1)\big) + (\xi_2-\zeta_2)\big(\psi'(\xi_2)-\psi'(\zeta_2)\big). \tag{5.14}$$

Therefore, Proposition 5.9 yields

$$H_A(x,u^*) \geqslant (\xi_1-\zeta_1)\big(\psi'(\xi_1)-\psi'(\zeta_1)\big) + (\xi_2-\zeta_2)\big(\psi'(\xi_2)-\psi'(\zeta_2)\big). \tag{5.15}$$

Let us note that the lower bound of Proposition 1.2(i) on the Haraux function would be harder to compute.

Next, we consider an application to the asymptotic behavior of a family of set-valued operators in terms of the warped resolvents of Definition 5.1.

Proposition 5.14. *Let $\mathfrak{A} = \big\{A_t\colon \mathcal{X} \to 2^{\mathcal{X}^*} \mid t \in \left[0,+\infty\right[\big\}$ be a family of operators, let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$ be ϕ-uniformly monotone, and let $\gamma \in \left]0,+\infty\right[$. Suppose that there exists an operator $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ such that*

$$(\forall (x,u^*) \in \operatorname{gra} A) \quad \lim_{t\to+\infty} H_{A_t}(x,u^*) = 0. \tag{5.16}$$

Additionally, suppose that, for every $B \in \mathfrak{A} \cup \{A\}$, $\mathcal{D} \cap \operatorname{dom} B \neq \varnothing$ and that one of properties (ii)–(ix) in Proposition 5.5 is satisfied, together with one of properties (i)–(iii) in Theorem 5.7. Then the following hold:

(i) *Let* $y^* \in \mathcal{X}^*$. *Then* $(W+\gamma A_t)^{-1}y^* \to (W+\gamma A)^{-1}y^*$.

(ii) *Let* $y \in \mathcal{D}$. *Then* $J^W_{\gamma A_t}y \to J^W_{\gamma A}y$.

(iii) [3, Proposition 2.1] *Suppose that* $\mathcal{X}$ *is Hilbertian, that* $W = \mathrm{Id}$, *and that the operators* $(A_t)_{t\in[0,+\infty[}$ *and* A *are maximally monotone. Then* $(A_t)_{t\in[0,+\infty[}$ *graph-converges to* A.

Proof. Set $x = (W+\gamma A)^{-1}y^*$, which is well defined since $W+\gamma A$ is surjective and injective. Note that $y^* \in Wx + \gamma Ax$ and set $u^* = \gamma^{-1}(y^* - Wx)$. Then $(x,u^*) \in \operatorname{gra} A$ and, in view of Proposition 5.2, Theorem 5.7 asserts that, for every $t \in [0,+\infty[$,

$$\begin{aligned} H_{A_t}(x,u^*) &\geqslant \frac{1}{\gamma}\left\langle x - J^W_{\gamma(A_t-u^*)}x, Wx - W\big(J^W_{\gamma(A_t-u^*)}x\big)\right\rangle \\ &= \frac{1}{\gamma}\left\langle x - (W+\gamma A_t)^{-1}(Wx+\gamma u^*), Wx - W\big((W+\gamma A_t)^{-1}(Wx+\gamma u^*)\big)\right\rangle \\ &\geqslant \frac{1}{\gamma}\phi\left(\left\| x - (W+\gamma A_t)^{-1}\left(Wx + \gamma\left(\frac{y^*-Wx}{\gamma}\right)\right)\right\|\right) \\ &= \frac{1}{\gamma}\phi\left(\left\|(W+\gamma A)^{-1}y^* - (W+\gamma A_t)^{-1}y^*\right\|\right). \end{aligned} \tag{5.17}$$

(i): This follows from (5.16) and (5.17).
(ii): Set $y^* = Wy$ in (5.17), and invoke (5.16) and Definition 5.1.
(iii): A direct consequence of (i) and [2, Proposition 3.60]. ☐

Outside of Hilbert spaces, the lower bounds on the Haraux function produced in this section via warped resolvents are different from those produced in Section 3 via metric resolvents. In particular, the lower bounds of Section 3 involve the distance to the graph of the operator. We show that it is possible to obtain such bounds with warped resolvents based on a certain type of kernel W. For this purpose, we introduce the following property, which implies the strong monotonicity of both W and its inverse.

Definition 5.15. Let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$, let $\alpha \in\,]0,+\infty[$, and let $\beta \in\,]0,+\infty[$. Then W is (α,β)-*jointly strongly monotone* if

$$(\forall x \in \mathcal{D})(\forall y \in \mathcal{D}) \quad \langle x-y, Wx-Wy\rangle \geqslant \frac{\alpha}{2}\|x-y\|^2 + \frac{\beta}{2}\|Wx-Wy\|^2. \tag{5.18}$$

Example 5.16. Let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$, let $\alpha \in\,]0,+\infty[$, and let $\beta \in\,]0,+\infty[$. Then the following hold:

(i) Suppose that W is α-strongly monotone. Then the following are satisfied:

  (a) Suppose that W is β-cocoercive, i.e.,

$$(\forall x \in \mathcal{D})(\forall y \in \mathcal{D}) \quad \langle x-y, Wx-Wy\rangle \geqslant \beta\|Wx-Wy\|^2. \tag{5.19}$$

  Then W is (α,β)-jointly strongly monotone.

  (b) Suppose that W is β-Lipschitzian. Then W is $(\alpha,\alpha/\beta^2)$-jointly strongly monotone.

  (c) Suppose that W is nonexpansive. Then W is (α,α)-jointly strongly monotone.

(ii) Let $f\colon \mathcal{X} \to \mathbb{R}$ be an α-strongly convex, Fréchet differentiable function. Suppose that ∇f is β^{-1}-Lipschitzian and that $W = \nabla f$. Then ∇f is (α,β)-jointly strongly monotone.

Proof. (i): Let $\{x, y\} \subset \mathcal{D}$. Since W is α-strongly monotone, we have $\langle x - y, Wx - Wy\rangle \geqslant \alpha\|x - y\|^2$.
(i)(a): Since W is β-cocoercive, we have $\langle x - y, Wx - Wy\rangle \geqslant \beta\|Wx - Wy\|^2$. Therefore,

$$\langle x - y, Wx - Wy\rangle \geqslant \max\{\alpha\|x - y\|^2, \beta\|Wx - Wy\|^2\} \geqslant \frac{\alpha}{2}\|x - y\|^2 + \frac{\beta}{2}\|Wx - Wy\|^2. \tag{5.20}$$

(i)(b): Since W is β-Lipschitzian, we get

$$\langle x - y, Wx - Wy\rangle \geqslant \frac{\alpha}{2}\|x - y\|^2 + \frac{\alpha}{2}\|x - y\|^2 \geqslant \frac{\alpha}{2}\|x - y\|^2 + \frac{\alpha}{2\beta^2}\|Wx - Wy\|^2. \tag{5.21}$$

(i)(c): Take $\beta = 1$ in (i)(b).
(ii): Since f is α-strongly convex, ∇f is α-strongly monotone. Additionally, ∇f is β-cocoercive by the Baillon–Haddad theorem [4, Corollaire 10]. We conclude by invoking (i)(a). ☐

Joint strong monotonicity allows us to derive from Proposition 5.3 lower bounds similar to those of Proposition 3.5.

Proposition 5.17. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$, let $x \in \mathcal{D}$, let $u^* \in \mathcal{X}^*$, let $\gamma \in \left]0, +\infty\right[$, let $\alpha \in \left]0, +\infty\right[$, and let $\beta \in \left]0, +\infty\right[$. Suppose that $z \in J^W_{\gamma(A-u^*)}x$ and that W is (α, β)-jointly strongly monotone. Then, using Notation 3.4,*

$$H_A(x, u^*) \geqslant \frac{\langle x - z, Wx - Wz\rangle}{\gamma} \geqslant \frac{\min\{\alpha, \beta\}}{2} d^2_{\mathrm{gra}\, A,\gamma}(x, u^*). \tag{5.22}$$

Proof. Proceeding as in (5.4), we observe that $(z, (Wx - Wz)/\gamma + u^*) \in \mathrm{gra}\, A$. It therefore follows from Proposition 5.3(iv), (5.18), and Notation 3.4 that

$$\begin{aligned} H_A(x, u^*) &\geqslant \frac{\langle x - z, Wx - Wz\rangle}{\gamma} \\ &\geqslant \frac{1}{\gamma}\left(\frac{\alpha}{2}\|x - z\|^2 + \frac{\beta}{2}\|Wx - Wz\|^2\right) \\ &\geqslant \frac{\min\{\alpha, \beta\}}{2}\left(\frac{\|x - z\|^2}{\gamma} + \gamma\left\|\frac{Wx - Wz}{\gamma} + u^* - u^*\right\|^2\right) \\ &\geqslant \frac{\min\{\alpha, \beta\}}{2} d^2_{\mathrm{gra}\, A,\gamma}(x, u^*), \end{aligned} \tag{5.23}$$

as claimed. ☐

Remark 5.18. When $\mathcal{X}$ is a Hilbert space, we deduce Proposition 3.5 from Proposition 5.17 applied with the kernel $W = \mathrm{Id}$, which is $(1, 1)$-jointly strongly monotone.

Theorem 5.19. *Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $\varnothing \neq \mathcal{D} \subset \mathcal{X}$, let $W\colon \mathcal{D} \to \mathcal{X}^*$, let $x \in \mathcal{D}$, let $u^* \in \mathcal{X}^*$, let $\gamma \in \left]0, +\infty\right[$, let $\alpha \in \left]0, +\infty\right[$, and let $\beta \in \left]0, +\infty\right[$. Suppose that $\mathcal{D} \cap \mathrm{dom}\, A \neq \varnothing$, that one of properties (i)–(ix) in Proposition 5.5 is satisfied, together with one of properties (i)–(iii) in Theorem 5.7, and that W is (α, β)-jointly strongly monotone. Set $z = J^W_{\gamma(A-u^*)}x$. Then, using Notation 3.4,*

$$H_A(x, u^*) \geqslant \frac{\langle x - z, Wx - Wz\rangle}{\gamma} \geqslant \frac{\min\{\alpha, \beta\}}{2} d^2_{\mathrm{gra}\, A,\gamma}(x, u^*). \tag{5.24}$$

Proof. This follows from Theorem 5.7 and Proposition 5.17. ☐

§6. Lower bounds on the Fenchel–Young function based on warped proximity operators

Following the pattern adopted in Section 4, we derive from the results of Section 5 lower bounds on the Fenchel–Young function of (1.4). Comparisons with existing bounds are made in several examples. To this end, we need the following extension of [15, Example 3.1] to set-valued kernels.

Definition 6.1. Let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be proper and let $K\colon \mathcal{X} \to 2^{\mathcal{X}^*}$. Then the *warped proximity operator* of φ with kernel K is $\operatorname{prox}_\varphi^K = J_{\partial\varphi}^K = (K+\partial\varphi)^{-1} \circ K$.

Here is a consequence of Lemma 4.1, Proposition 5.3, and Definition 6.1.

Proposition 6.2. *Let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be proper, let $W\colon \mathcal{X} \to 2^{\mathcal{X}^*}$, let $u^* \in \mathcal{X}^*$, and let $\gamma \in \left]0,+\infty\right[$. Suppose that $x \in \operatorname{dom} W$ and that $z \in \operatorname{prox}_{\gamma(\varphi-u^*)}^W x$. Then the following hold:*

(i) *There exist $x^* \in Wx$ and $z^* \in Wz$ such that $L_\varphi(x,u^*) \geqslant \langle x-z, x^*-z^*\rangle/\gamma$.*

(ii) *Suppose that W is ϕ-uniformly monotone. Then $L_\varphi(x,u^*) \geqslant \phi(\|x-z\|)/\gamma$.*

(iii) *Suppose that W is α-strongly monotone. Then $L_\varphi(x,u^*) \geqslant \alpha\|x-z\|^2/\gamma$.*

(iv) *Suppose that W is at most single-valued. Then $L_\varphi(x,u^*) \geqslant \langle x-z, Wx-Wz\rangle/\gamma$.*

(v) *Let $f \in \Gamma_0(\mathcal{X})$ be Gateaux differentiable on $\operatorname{dom}\nabla f = \operatorname{int}\operatorname{dom} f$ and suppose that $W = \nabla f$. Then*

$$L_\varphi(x,u^*) \geqslant \frac{D_f(x,z)+D_f(z,x)}{\gamma}. \tag{6.1}$$

Remark 6.3. We can also derive Proposition 6.2 directly, without invoking Proposition 5.3. Indeed, by Definition 6.1, since $z \in \operatorname{prox}_{\gamma(\varphi-u^*)}^W x$, there exists $x^* \in Wx$ such that $x^* \in Wz + \partial(\gamma(\varphi-u^*))(z) = Wz + \gamma\partial\varphi(z) - \gamma u^*$. Hence, there exists $z^* \in Wz$ such that $\gamma^{-1}(x^*-z^*)+u^* \in \partial\varphi(z)$. In turn, (2.5) yields

$$(\forall w \in \mathcal{X}) \quad \left\langle w-z, \frac{x^*-z^*}{\gamma}+u^*\right\rangle + \varphi(z) \leqslant \varphi(w). \tag{6.2}$$

Thus, $\langle x-z, x^*-z^*\rangle/\gamma + \langle x-z,u^*\rangle + \varphi(z) \leqslant \varphi(x)$ and, since $z \in \operatorname{dom}\varphi$, we conclude that

$$\frac{\langle x-z, x^*-z^*\rangle}{\gamma} \leqslant \varphi(x) - \langle x,u^*\rangle + \langle z,u^*\rangle - \varphi(z) \leqslant \varphi(x) - \langle x,u^*\rangle + \varphi^*(u^*) = L_\varphi(x,u^*). \tag{6.3}$$

This shows (i) in Proposition 6.2. Items (ii)–(v) in Proposition 6.2 then follow as in Proposition 5.3.

Next, we characterize the situation in which z in Proposition 6.2(iv) is equal to x.

Proposition 6.4. *In the setting of Proposition 6.2(iv), consider the following statements:*

[a] $x \in \operatorname{prox}_{\gamma(\varphi-u^*)}^W x$.

[b] $u^* \in \partial\varphi(x)$.

[c] $L_\varphi(x,u^*) = 0$.

[d] $H_{\partial\varphi}(x,u^*) = 0$.

Then the following hold:

(i) [a]⇔[b]⇔[c]⇒[d].

(ii) *Suppose that* $\varphi \in \Gamma_0(\mathcal{X})$*. Then* [c]$\Leftrightarrow$[d].

Proof. We recall that $\partial\varphi$ is monotone, and maximally so if $\varphi \in \Gamma_0(\mathcal{X})$. The claims therefore follow from (1.5) and Proposition 5.4. ☐

Let $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ be proper. Setting $A = \partial\varphi$ in Theorem 5.7 furnishes a first set of conditions under which z exists in Proposition 6.2 and is uniquely defined as $z = \operatorname{prox}^W_{\gamma(\varphi-u^*)} x$. Below, we refine some of these conditions and add new ones.

Theorem 6.5. *Let* $\varphi\colon \mathcal{X} \to \left]-\infty,+\infty\right]$ *be proper, let* $\varnothing \neq \mathcal{D} \subset \mathcal{X}$*, let* $W\colon \mathcal{D} \to \mathcal{X}^*$*, let* $x \in \mathcal{D}$*,* $u^* \in \mathcal{X}^*$*, and let* $\gamma \in \left]0,+\infty\right[$*. Suppose that* $\mathcal{D} \cap \operatorname{dom}\partial\varphi \neq \varnothing$ *and that one of properties* (i)–(iii) *below is satisfied, together with one of properties* (iv)–(vii)*:*

(i) $W + \gamma\partial\varphi$ *is injective.*

(ii) $W + \gamma\partial\varphi$ *is strictly monotone.*

(iii) W *is uniformly monotone.*

(iv) $Wx + \gamma u^* \in \operatorname{ran}(W + \gamma\partial\varphi)$.

(v) W *is monotone,* $W + \gamma\partial\varphi$ *is maximally monotone,* $\operatorname{ran} W + \gamma \operatorname{ran}\partial\varphi = \mathcal{X}^*$*, and one of the following is satisfied:*

  (a) W *is* 3^* *monotone.*

  (b) $\operatorname{dom} W \subset \operatorname{dom}\partial\varphi$.

(vi) $W = \partial f$*, with* $f \in \Gamma_0(\mathcal{X})$*,* $\varphi \in \Gamma_0(\mathcal{X})$*, the cone generated by* $\operatorname{dom} f - \operatorname{dom}\varphi$ *is a closed vector subspace of* $\mathcal{X}$*, and* $\operatorname{dom}\partial f^* + \gamma \operatorname{dom}\partial\varphi^* = \mathcal{X}^*$.

(vii) W *is monotone,* $W + \gamma\partial\varphi$ *is maximally monotone, and one of the following is satisfied:*

  (a) W *is* 3^* *monotone and surjective.*

  (b) $\partial\varphi$ *is surjective.*

Set $z = \operatorname{prox}^W_{\gamma(\varphi-u^*)} x$*. Then* $L_\varphi(x,u^*) \geqslant \langle x - z, Wx - Wz\rangle/\gamma$.

Proof. We apply Theorem 5.7 with $A = \partial\varphi$, taking into account the fact that, by [41, Proposition 32.42],

$$A \text{ is } 3^* \text{ monotone.} \tag{6.4}$$

(i): Theorem 5.7(i).
(ii): Theorem 5.7(ii).
(iii): Theorem 5.7(iii) and (6.4).
(iv): Proposition 5.5(i).
(v): Proposition 5.5(vii) and (6.4).
(vi): This follows from Proposition 5.5(viii) and (6.4). Indeed, by the Attouch–Brézis theorem [38, Theorem 18.2], $W + \gamma A = \partial f + \gamma\partial\varphi = \partial(f + \gamma\varphi)$. However, since $f + \gamma\varphi \in \Gamma_0(\mathcal{X})$, Rockafellar's theorem [38, Theorem 18.7] asserts that this operator is maximally monotone.
(vii): Proposition 5.5(ix) and (6.4). ☐

We now focus on the case when W is the gradient of a Legendre function f. The warped proximity operator $\operatorname{prox}^{\nabla f}_\varphi$ becomes the Bregman proximity operator of φ studied in [8, Section 3.4].

Proposition 6.6. *Let* $\varphi \in \Gamma_0(\mathcal{X})$*, let* $f \in \Gamma_0(\mathcal{X})$ *be a Legendre function, let* $x \in \operatorname{int}\operatorname{dom} f$*, let* $u^* \in \mathcal{X}^*$*, and let* $\gamma \in \left]0,+\infty\right[$*. Suppose that* $(\operatorname{int}\operatorname{dom} f) \cap \operatorname{dom}\partial\varphi \neq \varnothing$ *and that one of the following holds:*

(i) $\gamma u^* + \operatorname{int}\operatorname{dom} f^* \subset \operatorname{ran}(\nabla f + \gamma\partial\varphi)$.

(ii) $(\operatorname{int}\operatorname{dom} f^*) + \gamma \operatorname{dom}\partial\varphi^* = \mathcal{X}^*$.

(iii) $f + \gamma\varphi$ *is cofinite.*

(iv) $f + \gamma\varphi$ *is supercoercive.*

(v) $\operatorname{dom} f \cap \operatorname{dom}\varphi$ *is bounded.*

Set $z = \operatorname{prox}^{\nabla f}_{\gamma(\varphi - u^*)} x$. *Then*

$$L_\varphi(x, u^*) \geqslant \frac{D_f(x, z) + D_f(z, x)}{\gamma}. \tag{6.5}$$

Proof. (i)–(ii): These follow from Proposition 5.9, (6.4), and Lemma 4.1.

(iii)⇒(i): We have $f + \gamma\varphi \in \Gamma_0(\mathcal{X})$, hence $(f + \gamma\varphi)^* \in \Gamma_0(\mathcal{X}^*)$. Therefore, by [40, Corollary 47.7, Theorem 47.A(ii), Theorem 51.A(ii), and Theorem 47.B] and [7, Theorem 5.6],

$$\begin{aligned} \mathcal{X}^* &= \operatorname{int}\operatorname{dom}(f + \gamma\varphi)^* \\ &\subset \operatorname{dom}\partial(f + \gamma\varphi)^* \\ &= \operatorname{ran}\big(\partial(f + \gamma\varphi)^*\big)^{-1} \\ &= \operatorname{ran}\partial(f + \gamma\varphi) \\ &= \operatorname{ran}(\partial f + \gamma\partial\varphi) \\ &= \operatorname{ran}(\nabla f + \gamma\partial\varphi), \end{aligned} \tag{6.6}$$

which confirms that $\operatorname{ran}(\nabla f + \gamma\partial\varphi) = \mathcal{X}^*$.

(iv)⇒(iii): [7, Theorem 3.4].

(v)⇒(iv): Clear. ☐

Example 6.7. When $\mathcal{X}$ is a Hilbert space, we recover Proposition 1.2(ii) as a special case of Proposition 6.6(ii) with $f = \|\cdot\|^2/2$. It is also a special case of Proposition 6.2(iv) with $W = \mathrm{Id}$.

Example 6.8. Let $\varphi \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $x \in \operatorname{int}\operatorname{dom}\varphi$, let $u^* \in \operatorname{int}\operatorname{dom}\varphi^*$, and let $\gamma \in \left]0, +\infty\right[$. Then

$$L_\varphi(x, u^*) \geqslant \frac{\Big\langle x - \nabla\varphi^*\Big((1+\gamma)^{-1}\big(\nabla\varphi(x) + \gamma u^*\big)\Big), \nabla\varphi(x) - u^*\Big\rangle}{1 + \gamma}. \tag{6.7}$$

Proof. It follows from the results of [7, Section 5] that $\operatorname{dom}\partial\varphi^* = \operatorname{int}\operatorname{dom}\varphi^*$ and $\nabla\varphi \colon \operatorname{int}\operatorname{dom}\varphi \to \operatorname{int}\operatorname{dom}\varphi^*$ is a bijection with inverse $\nabla\varphi^*$. We establish the claim by setting $f = \varphi$ in Proposition 6.6(i). We first observe that

$$\gamma u^* + \operatorname{int}\operatorname{dom}\varphi^* \subset (\gamma + 1)\operatorname{int}\operatorname{dom}\varphi^* = (1+\gamma)\operatorname{dom}\partial\varphi^* = (1+\gamma)\operatorname{ran}\partial\varphi = \operatorname{ran}\big((1+\gamma)\partial\varphi\big). \tag{6.8}$$

As in Proposition 6.6(i), set $z = \operatorname{prox}^{\nabla\varphi}_{\gamma(\varphi - u^*)} x = ((1+\gamma)\nabla\varphi - \gamma u^*)^{-1}(\nabla\varphi(x))$. Then $\nabla\varphi(x) + \gamma u^* =$

$(1+\gamma)\nabla\varphi(z)$ and therefore $z = \nabla\varphi^*((1+\gamma)^{-1}(\nabla\varphi(x)+\gamma u^*))$. Thus,

$$\begin{aligned}
\frac{\big\langle x - z, \nabla\varphi(x) - \nabla\varphi(z)\big\rangle}{\gamma} &= \frac{\Big\langle x - z, \nabla\varphi(x) - (1+\gamma)^{-1}\big(\nabla\varphi(x)+\gamma u^*\big)\Big\rangle}{\gamma} \\
&= \frac{\Big\langle x - z, \gamma\big(\nabla\varphi(x) - u^*\big)\Big\rangle}{\gamma(1+\gamma)} \\
&= \frac{\Big\langle x - \nabla\varphi^*\Big((1+\gamma)^{-1}\big(\nabla\varphi(x)+\gamma u^*\big)\Big), \nabla\varphi(x) - u^*\Big\rangle}{1+\gamma}
\end{aligned} \tag{6.9}$$

which, in view of (6.5) and (2.6), yields (6.7). ☐

Example 6.9. Suppose that $\mathcal{X}$ is a Hilbert space with scalar product $\langle\cdot\,|\,\cdot\rangle$, let $\psi \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $\widetilde{\psi}\colon \mathcal{X} \to \mathbb{R}\colon x \mapsto \inf_{y\in\mathcal{X}}(\psi(y)+\|x-y\|^2/2)$ be its Moreau envelope, and set $\varphi = \|\cdot\|^2/2+\psi$. Let $x \in \operatorname{int}\operatorname{dom}\psi$, let $u^* \in \mathcal{X}$, and let $\gamma \in \,]0,+\infty[$. Then

$$L_\varphi(x,u^*) = \frac{1}{2}\|x-u^*\|^2 + \psi(x) - \widetilde{\psi}(u^*). \tag{6.10}$$

Further, Proposition 1.2(ii) gives

$$L_\varphi(x,u^*) \geqslant \frac{\left\|x - \operatorname{prox}_{\gamma(1+\gamma)^{-1}\psi}\left(\dfrac{x+\gamma u^*}{1+\gamma}\right)\right\|^2}{\gamma}, \tag{6.11}$$

while Example 6.8 gives

$$L_\varphi(x,u^*) \geqslant \frac{\langle x - z\,|\,x + \nabla\psi(x) - u^*\rangle}{1+\gamma}, \quad \text{where} \quad z = \operatorname{prox}_\psi\left(\frac{x+\nabla\psi(x)+\gamma u^*}{1+\gamma}\right). \tag{6.12}$$

Proof. By [9, Example 13.4], $\varphi^* = \|\cdot\|^2/2 - \widetilde{\psi}$. Therefore,

$$L_\varphi(x,u^*) = \frac{1}{2}\|u^*\|^2 + \frac{1}{2}\|x\|^2 - \langle x\,|\,u^*\rangle + \psi(x) - \widetilde{\psi}(u^*) = \frac{1}{2}\|x-u^*\|^2 + \psi(x) - \widetilde{\psi}(u^*). \tag{6.13}$$

We derive (6.11) from Proposition 1.2(ii) and [9, Proposition 24.8(i)]. We also note that φ is a Legendre function. Moreover, $\operatorname{dom}\varphi^* = \mathcal{X}$ and $\nabla\varphi^* = \operatorname{prox}_\psi$ [9, Proposition 12.30]. In turn, we derive (6.12) from Example 6.8. ☐

Remark 6.10. In Example 6.9, set $\psi = \|\cdot\|^2/2$ and $\gamma = 1$. Then $L_\varphi(x,u^*) = \|2x-u^*\|^2/4$. Furthermore, Proposition 1.2(ii) yields

$$L_\varphi(x,u^*) \geqslant \left\|x - \operatorname{prox}_{\|\cdot\|^2/4}\left(\frac{x+u^*}{2}\right)\right\|^2 = \frac{\|2x-u^*\|^2}{9}. \tag{6.14}$$

On the other hand, since $\nabla\psi = \mathrm{Id}$ in (6.12), we have

$$z = \operatorname{prox}_\psi\left(\frac{2x+u^*}{2}\right) = \frac{2x+u^*}{4} \tag{6.15}$$

and therefore the minorization of Example 6.8 becomes

$$L_\varphi(x,u^*) \geqslant \frac{\langle x-z\,|\,2x-u^*\rangle}{2} = \frac{\|2x-u^*\|^2}{8}. \tag{6.16}$$

The lower bound produced by Example 6.8 is therefore sharper than that of Proposition 1.2(ii).

Our lower bounds on the Haraux and Fenchel–Young functions are new, even in Euclidean spaces. Here are two examples in which they are compared to the lower bound of Proposition 1.2(ii).

Example 6.11. Let $\mathcal{X}$ be the standard Euclidean space $\mathbb{R}^N$ and $I = \{1, \ldots, N\}$. Consider the negative Burg entropy function

$$\varphi\colon \mathcal{X} \to \left]-\infty, +\infty\right] : x = (\xi_i)_{i\in I} \mapsto \begin{cases} -\sum_{i\in I} \ln(\xi_i), & \text{if } x \in \left]0, +\infty\right[^N; \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.17}$$

Let $x \in \left]0, +\infty\right[^N$, $u^* = (\mu_i^*)_{i\in I} \in \left]-\infty, 0\right[^N$, and $\gamma \in \left]0, +\infty\right[$. Then

$$L_\varphi(x, u^*) = -N - \sum_{i\in I}\big(\ln(-\xi_i\mu_i^*) + \xi_i\mu_i^*\big) \tag{6.18}$$

and Example 6.8 gives

$$L_\varphi(x, u^*) \geqslant \sum_{i\in I} \frac{\gamma|1 + \xi_i\mu_i^*|^2}{(1+\gamma)(1-\gamma\xi_i\mu_i^*)}. \tag{6.19}$$

Let us observe that, if $N = 1$, this lower bound becomes $\gamma|1 + \xi_1\mu_1^*|^2/((1+\gamma)(1-\gamma\xi_1\mu_1^*))$. In comparison, we derive from [9, Example 24.40] that the lower bound given by Proposition 1.2(ii) is

$$\frac{\left|\xi_1 - \gamma\mu_1^* - \sqrt{|\xi_1 + \gamma\mu_1^*|^2 + 4\gamma}\right|^2}{4\gamma}. \tag{6.20}$$

We graph these two bounds in Figure 1 for different values of γ, which shows that the bound provided by Proposition 1.2(ii) is not the best.

Proof. By [9, Example 13.2(iii) and Proposition 13.30], $\varphi^*(u^*) = -N - \sum_{i\in I}\ln(-\mu_i^*)$. Thus,

$$L_\varphi(x, u^*) = -N - \sum_{i\in I}\big(\ln(-\xi_i\mu_i^*) + \xi_i\mu_i^*\big). \tag{6.21}$$

Remark that φ is a Legendre function with $\operatorname{dom}\nabla\varphi = \operatorname{int}\operatorname{dom}\varphi = \left]0, +\infty\right[^N$ and

$$\big(\forall y = (\eta_i)_{i\in I} \in \left]0, +\infty\right[^N\big) \quad \nabla\varphi(y) = \big(-1/\eta_i\big)_{i\in I}. \tag{6.22}$$

Further, it is clear that $u^* \in \left]-\infty, 0\right[^N = \operatorname{int}\operatorname{dom}\varphi^*$. By Example 6.8,

$$z = (\zeta_i)_{i\in I} = \nabla\varphi^*\Big((1+\gamma)^{-1}\big(\nabla\varphi(x) + \gamma u^*\big)\Big) \tag{6.23}$$

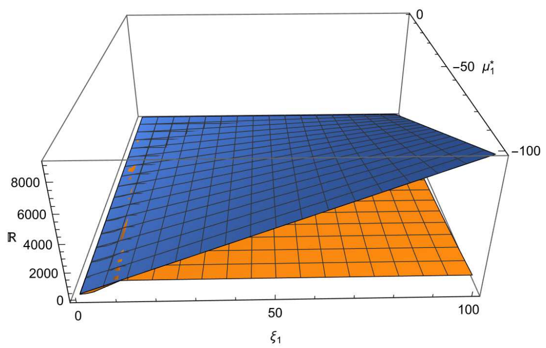
is well defined and $\nabla\varphi(x) + \gamma u^* = (1+\gamma)\nabla\varphi(z)$, which yields $(\forall i \in I)$ $(1 - \gamma\xi_i\mu_i^*)/\xi_i = (1+\gamma)/\zeta_i$. Hence,

$$z = \left(\frac{(1+\gamma)\xi_i}{1 - \gamma\xi_i\mu_i^*}\right)_{i\in I} \in \left]0, +\infty\right[^N. \tag{6.24}$$

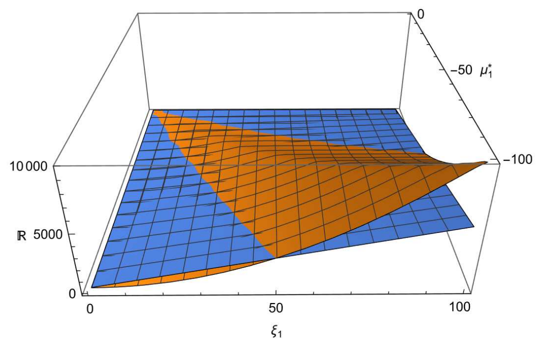
We thus derive from (2.6) that

$$D_\varphi(x, z) + D_\varphi(z, x) = \sum_{i\in I}\left(-2 + \frac{\xi_i}{\zeta_i} + \frac{\zeta_i}{\xi_i}\right) = \sum_{i\in I} \frac{\gamma|1 + \xi_i\mu_i^*|^2}{(1+\gamma)(1-\gamma\xi_i\mu_i^*)}, \tag{6.25}$$

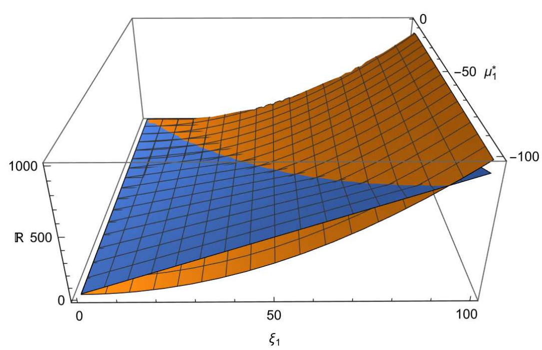
which concludes the proof. ☐



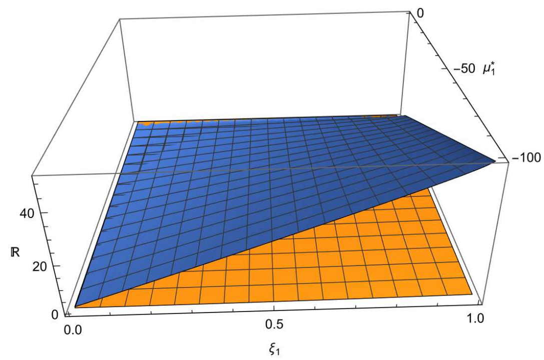
(a) Example 6.11, $\gamma = 0.1$.



(b) Example 6.11, $\gamma = 1$.



(c) Example 6.11, $\gamma = 10$.



(d) Example 6.12, $\gamma = 1$.

Figure 1: Lower bounds of Examples 6.11 and 6.12. In blue, the lower bound of Proposition 6.6 and, in orange, the lower bound of Proposition 1.2(ii).

Example 6.12. Let $\mathcal{X}$ be the standard Euclidean space $\mathbb{R}^N$ and $I = \{1, \ldots, N\}$. Set

$$\begin{aligned}(\forall x = (\xi_i)_{i\in I} \in \mathcal{X}) \quad I_1(x) = \big\{i \in I \mid \xi_i \in [0, +\infty[\big\},\ \ I_2(x) = \big\{i \in I \mid \xi_i \in \,]0, +\infty[\big\},\\ I_3(x) = \big\{i \in I \mid \xi_i \in [0, 1]\big\},\ \text{ and }\ I_4(x) = \big\{i \in I \mid \xi_i \in \,]0, 1[\big\}.\end{aligned} \tag{6.26}$$

We consider the negative Boltzmann–Shannon entropy function

$$\varphi\colon \mathcal{X} \to \,]-\infty, +\infty] : x \mapsto \begin{cases} \displaystyle\sum_{i\in I_2(x)} \big(\xi_i \ln(\xi_i) - \xi_i\big), & \text{if } I = I_1(x) \text{ and } I_2(x) \neq \varnothing;\\ 0, & \text{if } I = I_1(x) \smallsetminus I_2(x);\\ +\infty, & \text{otherwise,}\end{cases} \tag{6.27}$$

and let f be the Fermi–Dirac entropy function

$$f\colon \mathcal{X} \to \,]-\infty, +\infty] : x \mapsto \begin{cases} \displaystyle\sum_{i\in I_4(x)} \big(\xi_i \ln(\xi_i) + (1-\xi_i)\ln(1-\xi_i)\big), & \text{if } I = I_3(x) \text{ and } I_4(x) \neq \varnothing;\\ 0, & \text{if } I = I_3(x) \smallsetminus I_4(x);\\ +\infty, & \text{otherwise.}\end{cases} \tag{6.28}$$

Let $x \in \left]0,1\right[^N$, $u^* = (\mu_i^*)_{i\in I} \in \mathcal{X}$, and $\gamma \in \left]0,+\infty\right[$. Set

$$(\forall i \in I) \quad \zeta_i = -\frac{\xi_i e^{\gamma\mu_i^*}}{2(1-\xi_i)} + \sqrt{\frac{\xi_i^2 e^{2\gamma\mu_i^*}}{4|1-\xi_i|^2} + \frac{\xi_i e^{\gamma\mu_i^*}}{1-\xi_i}}. \tag{6.29}$$

Then

$$L_\varphi(x,u^*) = \sum_{i\in I}\big(\xi_i \ln(\xi_i) - \xi_i + e^{\mu_i^*} - \xi_i\mu_i^*\big) \tag{6.30}$$

and Proposition 6.6(v) gives

$$L_\varphi(x,u^*) \geqslant \frac{1}{\gamma}\sum_{i\in I}(\xi_i - \zeta_i)\ln\left(\frac{\xi_i(1-\zeta_i)}{\zeta_i(1-\xi_i)}\right). \tag{6.31}$$

Thus, if $N = 1$ and $\gamma = 1$, the above lower bound is

$$(\xi_1 - \zeta_1)\ln\left(\frac{\xi_1(1-\zeta_1)}{\zeta_1(1-\xi_1)}\right). \tag{6.32}$$

By contrast, since $\operatorname{prox}_\varphi \colon \xi_1 \mapsto \varLambda(e^{\xi_1})$, where $\varLambda$ is the Lambert W-function [9, Example 24.39], the lower bound of Proposition 1.2(ii) is

$$\big|\xi_1 - \varLambda\big(e^{\xi_1+\mu_1^*}\big)\big|^2. \tag{6.33}$$

We illustrate these bounds in Figure 1, which shows that the bound given by Proposition 1.2(ii) is not the best. The bound (6.32) provided by Proposition 6.6 is also easier to compute.

Proof. Using [9, Example 13.2(v) and Proposition 13.30], we obtain $\varphi^*(u^*) = \sum_{i\in I} e^{\mu_i^*}$ and hence

$$L_\varphi(x,u^*) = \sum_{i\in I}\big(\xi_i \ln(\xi_i) - \xi_i + e^{\mu_i^*} - \xi_i\mu_i^*\big). \tag{6.34}$$

Further, f is a Legendre function [6, Sections 5 and 6], with $\operatorname{dom}\nabla f = \operatorname{int}\operatorname{dom} f = \left]0,1\right[^N$, where

$$\Big(\forall y = (\eta_i)_{i\in I} \in \left]0,1\right[^N\Big) \quad \nabla f(y) = \big(\ln(\eta_i) - \ln(1-\eta_i)\big)_{i\in I}. \tag{6.35}$$

Additionally, $(\operatorname{int}\operatorname{dom} f) \cap \operatorname{dom}\partial\varphi = \left]0,1\right[^N \neq \varnothing$. Thus, by Proposition 6.6(v) and [25, Example 4.4], $z = (\zeta_i)_{i\in I}$ exists and

$$\begin{aligned}(\forall i \in I) \quad \zeta_i &= -\frac{e^{\ln(\xi_i)-\ln(1-\xi_i)+\gamma\mu_i^*}}{2} + \sqrt{\frac{e^{2\left(\ln(\xi_i)-\ln(1-\xi_i)+\gamma\mu_i^*\right)}}{4} + e^{\ln(\xi_i)-\ln(1-\xi_i)+\gamma\mu_i^*}} \\ &= -\frac{\xi_i e^{\gamma\mu_i^*}}{2(1-\xi_i)} + \sqrt{\frac{\xi_i^2}{4|1-\xi_i|^2}e^{2\gamma\mu_i^*} + \frac{\xi_i}{1-\xi_i}e^{\gamma\mu_i^*}}.\end{aligned} \tag{6.36}$$

Hence,

$$D_f(x,z) + D_f(z,x) = \sum_{i\in I}(\xi_i - \zeta_i)\ln\left(\frac{\xi_i(1-\zeta_i)}{\zeta_i(1-\xi_i)}\right) \tag{6.37}$$

and the conclusion follows. ☐

§7. Applications to monotone inclusions

As discussed in Section 1, finding lower bounds on the Haraux function — and hence on the Fenchel–Young function via Lemma 4.1 — is of both theoretical and practical interest. Thus, applications to inverse problems are discussed in [1], applications to the asymptotic properties of families of set-valued operators in [3, 34] and Proposition 5.14, applications to machine learning in [11, 35], applications to the strong Fitzpatrick inequality in [18], applications to convex analysis and convex programming in [20], and applications to optimal transportation in [20, 21]. In this section, we propose to apply the bounds of Sections 5 and 6 to the area of composite monotone inclusion problems.

Let $A\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be a maximally monotone operator. The Haraux function H_A of (1.1) is defined on the primal-dual space $\mathcal{X}\times\mathcal{X}^*$. We can employ it to induce a primal-primal function on $\mathcal{X}\times\mathcal{X}$. To this end, let $B\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ be maximally monotone and let $S\colon \operatorname{dom} B \to \mathcal{X}^*\colon x \mapsto Sx \in Bx$ be a selection of B. We associate with H_A the function

$$G_{A,S}\colon \mathcal{X}\times\mathcal{X} \to [0,+\infty]\colon (x,y) \mapsto H_A(x,-Sy). \tag{7.1}$$

This function can be used as a penalty function to detect whether a point $(x,y) \in \mathcal{X}\times\mathcal{X}$ is a zero of the direct sum operator $A \oplus B$ since Proposition 5.4 yields

$$G_{A,S}(x,y) = 0 \quad\Leftrightarrow\quad -Sy \in Ax \quad\Leftrightarrow\quad (x,y) \in \operatorname{zer}(A\oplus B). \tag{7.2}$$

As H_A, and therefore $G_{A,S}$, can be hard to evaluate, our results provide more tractable lower bounds to test the violation of the constraint $(x,y) \in \operatorname{zer}(A\oplus B)$. We can further specialize this construction to address the inclusion $0 \in Ax + Bx$, a generic model which covers a wide range of applications; see [24] and its bibliography. We introduce the primal function

$$\vartheta_{A,S}\colon \mathcal{X} \to [0,+\infty]\colon x \mapsto G_{A,S}(x,x) = H_A(x,-Sx) \tag{7.3}$$

to gauge the membership of a point $x \in \mathcal{X}$ in $\operatorname{zer}(A+B)$. Indeed, (7.2) yields

$$\vartheta_{A,S}(x) = 0 \quad\Leftrightarrow\quad -Sx \in Ax \quad\Leftrightarrow\quad x \in \operatorname{zer}(A+B). \tag{7.4}$$

To illustrate our lower bounds in this context, it is assumed henceforth that $B\colon \mathcal{X} \to \mathcal{X}^*$ is single-valued.

In the setting of Theorem 5.7, let $x \in \mathcal{D}$, set $u^* = -Bx$, and consider the kernel $K = W - \gamma B$. Then Proposition 5.2 yields

$$\begin{aligned} z &= J^{W}_{\gamma(A-u^*)}x \\ &= (W+\gamma A)^{-1}(Wx+\gamma u^*) \\ &= \big(W-\gamma B+\gamma(A+B)\big)^{-1}(Wx-\gamma Bx) \\ &= J^{K}_{\gamma(A+B)}x, \end{aligned} \tag{7.5}$$

and the lower bound of Theorem 5.7 is therefore

$$\vartheta_{A,B}(x) = H_A(x,-Bx) \geqslant \frac{\big\langle x - J^K_{\gamma(A+B)}x, Wx - W\big(J^K_{\gamma(A+B)}x\big)\big\rangle}{\gamma}. \tag{7.6}$$

Example 7.1. In the setting of Example 5.8, (7.6) yields

$$\vartheta_{A,B}(x) \geqslant \frac{\phi\Big(\big\|x - J^K_{\gamma(A+B)}x\big\|\Big)}{\gamma}, \quad \text{where} \quad K = W - \gamma B. \tag{7.7}$$

Example 7.2. In the setting of Proposition 5.9, (7.6) yields

$$\vartheta_{A,B}(x) \geqslant \frac{D_f\Big(x, J^K_{\gamma(A+B)}x\Big) + D_f\Big(J^K_{\gamma(A+B)}x, x\Big)}{\gamma}, \quad \text{where} \quad K = \nabla f - \gamma B. \tag{7.8}$$

In this case, $J^K_{\gamma(A+B)} = (\nabla f + \gamma A)^{-1} \circ (\nabla f - \gamma B)$ is the Bregman forward-backward splitting operator studied in [16].

Example 7.3. If $\mathcal{X}$ is Hilbertian and $f = \|\cdot\|^2/2$ in Example 7.2, then (7.8) becomes

$$\vartheta_{A,B}(x) \geqslant \frac{\big\|x - J_{\gamma A}(x - \gamma Bx)\big\|^2}{\gamma}, \tag{7.9}$$

which effectively splits A and B and is computable in terms of the standard resolvent $J_{\gamma A}$.

Example 7.4. We consider a primal-dual composite problem discussed in [25]. Let $\mathcal{Y} \neq \{0\}$ be a reflexive real Banach space, let $\boldsymbol{\mathcal{E}}$ be the standard product vector space $\mathcal{X} \times \mathcal{Y}^*$ equipped with the norm $(x, y^*) \mapsto \sqrt{\|x\|^2 + \|y^*\|^2}$, and let $\boldsymbol{\mathcal{E}}^*$ be its topological dual, that is, $\mathcal{X}^* \times \mathcal{Y}$ equipped with the norm $(x^*, y) \mapsto \sqrt{\|x^*\|^2 + \|y\|^2}$. Let $C\colon \mathcal{X} \to 2^{\mathcal{X}^*}$ and $D\colon \mathcal{Y} \to 2^{\mathcal{Y}^*}$ be maximally monotone, and let $L\colon \mathcal{X} \to \mathcal{Y}$ be linear and bounded. Under consideration is the primal-dual system

$$\text{find} \;\; (x, y^*) \in \boldsymbol{\mathcal{E}} \;\; \text{such that} \;\; \begin{cases} 0 \in Cx + L^*\big(D(Lx)\big) \\ 0 \in -L\big(C^{-1}(-L^*y^*)\big) + D^{-1}y^*. \end{cases} \tag{7.10}$$

Let us introduce the operators

$$\begin{cases} \boldsymbol{A}\colon \boldsymbol{\mathcal{E}} \to 2^{\boldsymbol{\mathcal{E}}^*}\colon (x, y^*) \mapsto Cx \times D^{-1}y^* \\ \boldsymbol{B}\colon \boldsymbol{\mathcal{E}} \to \boldsymbol{\mathcal{E}}^*\colon (x, y^*) \mapsto (L^*y^*, -Lx). \end{cases} \tag{7.11}$$

As shown in [25, Section 2.1], $\boldsymbol{A}$ and $\boldsymbol{B}$ are maximally monotone and every point in the Kuhn–Tucker set $\operatorname{zer}(\boldsymbol{A} + \boldsymbol{B})$ solves (7.10). Let $\gamma \in\,]0, +\infty[$, $\mathcal{D}_{\mathcal{X}} \subset \mathcal{X}$, and $\mathcal{D}_{\mathcal{Y}^*} \subset \mathcal{Y}^*$. Let $W_{\mathcal{X}}\colon \mathcal{D}_{\mathcal{X}} \to \mathcal{X}^*$ and $W_{\mathcal{Y}^*}\colon \mathcal{D}_{\mathcal{Y}^*} \to \mathcal{Y}$ be such that $W_{\mathcal{X}} + \gamma C$ and $W_{\mathcal{Y}^*} + \gamma D^{-1}$ are surjective and injective. Further, set $\boldsymbol{\mathcal{D}} = \mathcal{D}_{\mathcal{X}} \times \mathcal{D}_{\mathcal{Y}^*}$ and $\boldsymbol{W}\colon \boldsymbol{\mathcal{D}} \to \boldsymbol{\mathcal{E}}^*\colon (x, y^*) \mapsto (W_{\mathcal{X}}x, W_{\mathcal{Y}^*}y^*)$. Then $\boldsymbol{W} + \gamma\boldsymbol{A}$ is surjective and injective, which confirms that properties (ii) of Proposition 5.5 and (i) of Theorem 5.7 are satisfied. Additionally, we define $\boldsymbol{K} = \boldsymbol{W} - \gamma\boldsymbol{B}\colon \boldsymbol{\mathcal{D}} \to \boldsymbol{\mathcal{E}}^*\colon (x, y^*) \mapsto (W_{\mathcal{X}}x - \gamma L^*y^*, W_{\mathcal{Y}^*}y^* + \gamma Lx)$. Now, let us fix $\boldsymbol{x} = (x, y^*) \in \boldsymbol{\mathcal{D}}$. Then, as in (7.4), how close $\boldsymbol{x}$ is to being a Kuhn–Tucker point can be gauged by the value of the penalty function $\vartheta(\boldsymbol{x}) = \vartheta_{\boldsymbol{A},\boldsymbol{B}}(\boldsymbol{x})$. Note that (7.5) and Proposition 5.2 entail that

$$\begin{aligned} J^{\boldsymbol{K}}_{\gamma(\boldsymbol{A}+\boldsymbol{B})}\boldsymbol{x} &= \big(\boldsymbol{K} + \gamma(\boldsymbol{A} + \boldsymbol{B})\big)^{-1}(\boldsymbol{K}\boldsymbol{x}) \\ &= \Big((W_{\mathcal{X}} + \gamma C)^{-1}(W_{\mathcal{X}}x - \gamma L^*y^*), (W_{\mathcal{Y}^*} + \gamma D^{-1})^{-1}(W_{\mathcal{Y}^*}y^* + \gamma Lx)\Big) \\ &= \Big(J^{W_{\mathcal{X}}}_{\gamma(C+L^*y^*)}x, J^{W_{\mathcal{Y}^*}}_{\gamma(D^{-1}-Lx)}y^*\Big) \end{aligned} \tag{7.12}$$

is well defined. Therefore, (7.6) applied to $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{K})$ in $\boldsymbol{\mathcal{E}}$ yields

$$\begin{aligned} \vartheta(\boldsymbol{x}) \geqslant \frac{1}{\gamma}\bigg(&\Big\langle x - J^{W_{\mathcal{X}}}_{\gamma(C+L^*y^*)}x, W_{\mathcal{X}}x - W_{\mathcal{X}}\Big(J^{W_{\mathcal{X}}}_{\gamma(C+L^*y^*)}x\Big)\Big\rangle \\ &+ \Big\langle y^* - J^{W_{\mathcal{Y}^*}}_{\gamma(D^{-1}-Lx)}y^*, W_{\mathcal{Y}^*}y^* - W_{\mathcal{Y}^*}\Big(J^{W_{\mathcal{Y}^*}}_{\gamma(D^{-1}-Lx)}y^*\Big)\Big\rangle\bigg). \end{aligned} \tag{7.13}$$

For example, suppose that $W_{\mathcal{X}}$ is maximally monotone, that the cone generated by $\mathcal{D}_{\mathcal{X}} - \operatorname{dom} C$ is a closed vector subspace of $\mathcal{X}$, and that $W_{\mathcal{X}}$ is $\phi_{\mathcal{X}}$-uniformly monotone with $\phi_{\mathcal{X}}(t)/t \to +\infty$ as $t \to +\infty$. Likewise, suppose that $W_{\mathcal{Y}^*}$ is maximally monotone, that the cone generated by $\mathcal{D}_{\mathcal{Y}^*} - \operatorname{ran} D$ is a closed vector subspace of $\mathcal{Y}^*$, and that $W_{\mathcal{Y}^*}$ is $\phi_{\mathcal{Y}^*}$-uniformly monotone with $\phi_{\mathcal{Y}^*}(t)/t \to +\infty$ as $t \to +\infty$. Then we deduce from (7.13) and Example 5.8 applied in $\mathcal{X}$ and in $\mathcal{Y}^*$, and from Proposition 5.2 that

$$\begin{aligned}\vartheta(\boldsymbol{x}) &\geqslant \frac{\phi_{\mathcal{X}}\Big(\big\|x - J^{W_{\mathcal{X}}}_{\gamma(C+L^*y^*)}x\big\|\Big) + \phi_{\mathcal{Y}^*}\Big(\big\|y^* - J^{W_{\mathcal{Y}^*}}_{\gamma(D^{-1}-Lx)}y^*\big\|\Big)}{\gamma}\\ &= \frac{\phi_{\mathcal{X}}\Big(\big\|x - (W_{\mathcal{X}} + \gamma C)^{-1}(W_{\mathcal{X}}x - \gamma L^*y^*)\big\|\Big) + \phi_{\mathcal{Y}^*}\Big(\big\|y^* - (W_{\mathcal{Y}^*} + \gamma D^{-1})^{-1}(W_{\mathcal{Y}^*}y^* + \gamma Lx)\big\|\Big)}{\gamma}.\end{aligned} \tag{7.14}$$

In particular, if $\mathcal{X}$ and $\mathcal{Y}$ are Hilbertian, $W_{\mathcal{X}} = \mathrm{Id}_{\mathcal{X}}$, and $W_{\mathcal{Y}^*} = \mathrm{Id}_{\mathcal{Y}^*}$, then all the above hypotheses are satisfied with $\phi_{\mathcal{X}} = \phi_{\mathcal{Y}^*} = |\cdot|^2$ and we obtain, for $\boldsymbol{x} = (x, y^*) \in \mathcal{X} \times \mathcal{Y}$,

$$\vartheta(\boldsymbol{x}) \geqslant \frac{\big\|x - J_{\gamma C}(x - \gamma L^*y^*)\big\|^2 + \big\|y^* - J_{\gamma D^{-1}}(y^* + \gamma Lx)\big\|^2}{\gamma}. \tag{7.15}$$

We conclude with an application of Example 7.4 to Fenchel–Rockafellar duality in optimization.

Example 7.5. Define $\mathcal{Y}$ and $\mathcal{E}$ as in Example 7.4, let $\varphi \in \Gamma_0(\mathcal{X})$, let $\psi \in \Gamma_0(\mathcal{Y})$, and let $L\colon \mathcal{X} \to \mathcal{Y}$ be linear and bounded. Consider the primal problem

$$\underset{x\in\mathcal{X}}{\text{minimize}}\;\; \varphi(x) + \psi(Lx) \tag{7.16}$$

and the dual problem

$$\underset{y^*\in\mathcal{Y}^*}{\text{minimize}}\;\; \varphi^*(-L^*y^*) + \psi^*(y^*). \tag{7.17}$$

Let us set $C = \partial\varphi$ and $D = \partial\psi$ in Example 7.4, and let us define $\boldsymbol{A}$ and $\boldsymbol{B}$ as in (7.11). Then every point in the Kuhn–Tucker set $\operatorname{zer}(\boldsymbol{A}+\boldsymbol{B})$ solves the primal-dual pair (7.16)–(7.17). Now let $\gamma \in \left]0, +\infty\right[$, and let $f \in \Gamma_0(\mathcal{X})$ and $g \in \Gamma_0(\mathcal{Y})$ be Legendre functions such that $\nabla f + \gamma\partial\varphi$ and $\nabla g^* + \gamma\partial\psi^*$ are surjective, and set $W_{\mathcal{X}} = \nabla f$ and $W_{\mathcal{Y}^*} = \nabla g^*$. We recall that g^* is a Legendre function [7, Corollary 5.5] and note that, since f and g^* are strictly convex on the convex sets $\mathcal{D}_{\mathcal{X}} = \operatorname{int}\operatorname{dom} f$ and $\mathcal{D}_{\mathcal{Y}^*} = \operatorname{int}\operatorname{dom} g^*$, respectively, $W_{\mathcal{X}}$ and $W_{\mathcal{Y}^*}$ are strictly monotone [41, Proposition 25.10]. So are therefore the operators $W_{\mathcal{X}} + \gamma C$ and $W_{\mathcal{Y}^*} + \gamma D^{-1}$ which, as in Theorem 6.5(ii), makes them injective. Next, we introduce $\mathrm{Prox}^f_{\gamma\varphi} = (W_{\mathcal{X}} + \gamma C)^{-1} = (\nabla f + \gamma\partial\varphi)^{-1}$ and $\mathrm{Prox}^{g^*}_{\gamma\psi^*} = (W_{\mathcal{Y}^*} + \gamma D^{-1})^{-1} = (\nabla g^* + \gamma\partial\psi^*)^{-1}$. Then, as in (7.8), we infer from (7.13) that, for every $\boldsymbol{x} = (x, y^*) \in \operatorname{int}\operatorname{dom} f \times \operatorname{int}\operatorname{dom} g^*$,

$$\begin{aligned}\vartheta(\boldsymbol{x}) \geqslant \frac{1}{\gamma}\Big(& D_f\Big(x, \mathrm{Prox}^f_{\gamma\varphi}\big(\nabla f(x) - \gamma L^*y^*\big)\Big) + D_f\Big(\mathrm{Prox}^f_{\gamma\varphi}\big(\nabla f(x) - \gamma L^*y^*\big), x\Big) +\\ & D_{g^*}\Big(y^*, \mathrm{Prox}^{g^*}_{\gamma\psi^*}\big(\nabla g^*(y^*) + \gamma Lx\big)\Big) + D_{g^*}\Big(\mathrm{Prox}^{g^*}_{\gamma\psi^*}\big(\nabla g^*(y^*) + \gamma Lx\big), y^*\Big)\Big).\end{aligned} \tag{7.18}$$

In particular, if $\mathcal{X}$ and $\mathcal{Y}$ are Hilbertian, $f = \|\cdot\|^2_{\mathcal{X}}/2$, $g = \|\cdot\|^2_{\mathcal{Y}}/2$, and $\boldsymbol{x} = (x, y^*) \in \mathcal{X} \times \mathcal{Y}$, we obtain

$$\vartheta(\boldsymbol{x}) \geqslant \frac{\big\|x - \mathrm{prox}_{\gamma\varphi}(x - \gamma L^*y^*)\big\|^2 + \big\|y^* - \mathrm{prox}_{\gamma\psi^*}(y^* + \gamma Lx)\big\|^2}{\gamma}, \tag{7.19}$$

which can also be viewed as a special case of (7.15). Note that, in the Hilbertian — and in particular the Euclidean — setting, $\operatorname{Prox}^{f}_{\gamma\varphi}$ and $\operatorname{Prox}^{g^*}_{\gamma\psi^*}$ may be easier to compute than $\operatorname{prox}_{\gamma\varphi}$ and $\operatorname{prox}_{\gamma\psi^*}$; see [5, 25, 31] and Example 6.12. This makes the lower bound of (7.18) more tractable than that of (7.19) in such instances.

§8. Conclusions and future directions

We have derived new lower bounds on the Haraux function of set-valued operators and on the Fenchel–Young function of proper functions in the framework of reflexive Banach spaces. These bounds were obtained by using two types of resolvents, namely metric resolvents and warped resolvents. Existing results have been recovered as special cases. Two directions for future work can be mentioned here:

- In Example 7.4, we have applied our results to primal-dual composite monotone inclusions of type (7.10). It appears that this approach can be used for general systems of multivariate inclusions involving more complex monotonicity-preserving operations such as those considered in [17] and [24, Section 10]. Indeed, as discussed in these papers, such primal-dual systems can be reformulated in "saddle form" and thus be reduced to solving a monotone inclusion of the type $\mathbf{0} \in \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}$ in a suitably defined product space $\mathbf{X}$, where $\mathbf{A}\colon \mathbf{X} \to 2^{\mathbf{X}^*}$ is maximally monotone and $\mathbf{B}\colon \mathbf{X} \to \mathbf{X}^*$ is cocoercive. We can then exploit (7.6) in this scenario.
- A question of interest is whether lower bounds on the Haraux function can be derived in the context of nonreflexive Banach spaces with maximally monotone operators of type (NI) [38, Section 36]. For instance, the lower bound (3.14) is known to hold in this context [39, Theorem 2.6] and, in the light of [39, Remark 2.4], it is reasonable to expect that the results of Sections 3 and 4 remain valid as well.

References

[1] F. Andrade, G. Peyré, and C. Poon, Learning from samples: Inverse problems over measures via sharpened Fenchel–Young losses, arXiv, 2025. https://arxiv.org/abs/2505.07124

[2] H. Attouch, *Variational Convergence for Functions and Operators.* Pitman, Boston, MA, 1984.

[3] H. Attouch, A. Cabot, and M.-O. Czarnecki, Asymptotic behavior of nonautonomous monotone and subgradient evolution equations, *Trans. Amer. Math. Soc.*, vol. 370, pp. 755–790, 2018.

[4] J.-B. Baillon and G. Haddad, Quelques propriétés des opérateurs angle-bornés et *n*-cycliquement monotones, *Israel J. Math.*, vol. 26, pp. 137–150, 1977.

[5] H. H. Bauschke, J. Bolte, and M. Teboulle, A descent lemma beyond Lipschitz gradient continuity: First-order methods revisited and applications, *Math. Oper. Res.*, vol. 42, pp. 330–348, 2017.

[6] H. H. Bauschke and J. M. Borwein, Legendre functions and the method of random Bregman projections, *J. Convex Anal.*, vol. 4, pp. 27–67, 1997.

[7] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.*, vol. 3, pp. 615–647, 2001.

[8] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, vol. 42, pp. 596–636, 2003.

[9] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Springer, New York, 2017.

[10] H. H. Bauschke, S. Singh, and X. Wang, On Carlier's inequality, *J. Convex Anal.*, vol. 30, pp. 499–514, 2023.

[11] M. Blondel, A. F. T. Martins, and V. Niculae, Learning with Fenchel–Young losses, *J. Machine Learn. Res.*, vol. 21, pp. 1–69, 2020.

[12] J. M. Borwein and J. D. Vanderwerff, *Convex Functions: Constructions, Characterizations and Counterexamples.* Cambridge University Press, Cambridge, 2010.

[13] H. Brézis and A. Haraux, Image d'une somme d'opérateurs monotones et applications, *Israel J. Math.*, vol. 23, pp. 165–186, 1976.

[14] M. N. Bùi, On a lemma by Brézis and Haraux, *J. Approx. Theory*, vol. 311, art. 106175, 2025.

[15] M. N. Bùi and P. L. Combettes, Warped proximal iterations for monotone inclusions, *J. Math. Anal. Appl.*, vol. 491, art. 124315, 2020.

[16] M. N. Bùi and P. L. Combettes, Bregman forward-backward operator splitting, *Set-Valued Var. Anal.*, vol. 29, pp. 583–603, 2021.

[17] M. N. Búi and P. L. Combettes, Multivariate monotone inclusions in saddle form, *Math. Oper. Res.*, vol. 47, pp. 1082–1109, 2022.

[18] R. S. Burachik and J. E. Martínez–Legaz, A note on Carlier's inequality, *Optimization*, published online 2025-10-09.

[19] R. S. Burachik and B. F. Svaiter, Maximal monotone operators, convex functions and a special family of enlargements, *Set-Valued Anal.*, vol. 10, pp. 297–316, 2002.

[20] G. Carlier, Fenchel–Young inequality with a remainder and applications to convex duality and optimal transport, *SIAM J. Optim.*, vol. 33, pp. 1463–1472, 2023.

[21] G. Carlier, P. Pegon, and L. Tamanini, Convergence rate of general entropic optimal transport costs, *Calc. Var.*, vol. 62, art. 116, 2023.

[22] L.-J. Chu, On the sum of monotone operators, *Michigan Math. J.*, vol. 43, pp. 273–289, 1996.

[23] I. Ciorănescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems.* Kluwer, Dordrecht, 1990.

[24] P. L. Combettes, The geometry of monotone operator splitting methods, *Acta Numer.*, vol. 33, pp. 487–632, 2024.

[25] P. L. Combettes and Q. V. Nguyen, Solving composite monotone inclusions in reflexive Banach spaces by constructing best Bregman approximations from their Kuhn–Tucker set, *J. Convex Anal.*, vol. 23, pp. 481–510, 2016.

[26] S. Fitzpatrick, Representing monotone operators by convex functions, in: *Workshop/Miniconference on Functional Analysis and Optimization.* Austral. Nat. Univ., vol. 20, pp. 59–65, 1988.

[27] A. Haraux, personal communication, August 2025.

[28] P.-J. Laurent, *Approximation et Optimisation.* Hermann, Paris, 1972.

[29] J. J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.*, vol. 255, pp. 2897–2899, 1962.

[30] J. J. Moreau, *Fonctionnelles Convexes* (Séminaire Jean Leray sur les Équations aux Dérivées Partielles, no. 2). Collège de France, Paris, 1966–1967.

[31] Q. V. Nguyen, Forward-backward splitting with Bregman distances, *Vietnam J. Math.*, vol. 45, pp. 519–539, 2017.

[32] T. Pennanen, On the range of monotone composite mappings, *J. Nonlinear Convex Anal.*, vol. 2, pp. 193–202, 2001.

[33] J.-P. Penot and R. Ratsimahalo, On the Yosida approximation of operators, *Proc. Roy. Soc. Edinburgh Sect. A*, vol. 131, pp. 945–966, 2001.

[34] J.-P. Penot and C. Zălinescu, On the convergence of maximal monotone operators, *Proc. Amer. Math. Soc.*, vol. 134, pp. 1937–1946, 2006.

[35] S. Rakotomandimby, J.-P. Chancelier, M. De Lara, and M. Blondel, Learning with Fitzpatrick losses, *Proc. 38th Conf. on Neural Inform. Process. Syst.*, pp. 79381–79409. Vancouver, Canada, December 10–15, 2024.

[36] S. Reich, The range of sums of accretive and monotone operators, *J. Math. Anal. Appl.*, vol. 68, pp. 310–317, 1979.

[37] S. Simons, *Minimax and Monotonicity.* Lecture Notes in Math. 1693, Springer, Berlin, 1998.

[38] S. Simons, *From Hahn–Banach to Monotonicity.* Lecture Notes in Math. 1693, Springer, Berlin, 2008.

[39] M. D. Voisei and C. Zălinescu, Strongly-representable monotone operators, *J. Convex Anal.*, vol. 16, pp. 1011–1033, 2009.

[40] E. Zeidler, *Nonlinear Functional Analysis and Its Applications III – Variational Methods and Optimization.* Springer, New York, 1985.

[41] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B – Nonlinear Monotone Operators.* Springer, New York, 1990.