Noname manuscript No.
(will be inserted by the editor)

Moreau's Decomposition in Banach Spaces

Patrick L. Combettes · **Noli N. Reyes**

Received: date / Accepted: date

Abstract Moreau's decomposition is a powerful nonlinear hilbertian analysis tool that has been used in various areas of optimization and applied mathematics. In this paper, it is extended to reflexive Banach spaces and in the context of generalized proximity measures. This extension unifies and significantly improves upon existing results.

Keywords Banach space · Bregman distance · convex optimization · infimal convolution · Legendre function · Moreau's decomposition · proximity operator

1 Introduction

Throughout this paper, $(\mathscr{X}, \|\cdot\|)$ is a reflexive real Banach space with topological dual $(\mathscr{X}^*, \|\cdot\|_*)$, and the canonical bilinear form on $\mathscr{X}\times\mathscr{X}^*$ is denoted by $\langle\cdot,\cdot\rangle$. The distance function to a set $C\subset\mathscr{X}$ is $d_C\colon x\mapsto \inf_{y\in C}\|x-y\|$, the metric projector onto C is $P_C\colon x\mapsto \big\{y\in C \;\big|\; \|x-y\| = d_C(x)\big\}$, and the polar cone of C is $C^{\ominus} = \big\{x^*\in\mathscr{X}^* \;\big|\; (\forall x\in C)\ \langle x, x^*\rangle\leq 0\big\}$. $\Gamma_0(\mathscr{X})$ is the class of lower semicontinuous convex functions $\varphi\colon \mathscr{X}\to\left]-\infty,+\infty\right]$ such that $\operatorname{dom}\varphi = \big\{x\in\mathscr{X} \;\big|\; \varphi(x)<+\infty\big\}\neq\varnothing$.

The work of P. L. Combettes was supported by the Agence Nationale de la Recherche under grant ANR-08-BLAN-0294-02.

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

Noli N. Reyes
University of the Philippines – Diliman
Institute of Mathematics
Quezon City, 1101 Philippines
E-mail: noli@math.upd.edu.ph

A classical tool in linear hilbertian analysis is the following orthogonal decomposition principle.

Proposition 1 *Suppose that $\mathscr{X}$ is a Hilbert space, let V be a closed vector subspace of $\mathscr{X}$ with orthogonal complement $V^\perp$, and let $x \in \mathscr{X}$. Then the following hold.*

(i) $\|x\|^2 = d_V^2(x) + d_{V^\perp}^2(x)$.
(ii) $x = P_V x + P_{V^\perp} x$.
(iii) $\langle P_V x, P_{V^\perp} x\rangle = 0$.

In 1962, Moreau proposed a nonlinear extension of this decomposition.

Proposition 2 [22] *Suppose that $\mathscr{X}$ is a Hilbert space, let K be a nonempty closed convex cone in $\mathscr{X}$, and let $x \in \mathscr{X}$. Then the following hold.*

(i) $\|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x)$.
(ii) $x = P_K x + P_{K^\ominus} x$.
(iii) $\langle P_K x, P_{K^\ominus} x\rangle = 0$.

Motivated by problems in unilateral mechanics, Moreau further extended this result in [23] (see also [25]). To state Moreau's decomposition principle, we require some basic notions from convex analysis [7, 33]. Let φ and f be two functions in $\Gamma_0(\mathscr{X})$. The conjugate of φ is the function φ^* in $\Gamma_0(\mathscr{X}^*)$ defined by

$$\varphi^* \colon \mathscr{X}^* \to \left]-\infty,+\infty\right] \colon x^* \mapsto \sup_{x\in\mathscr{X}} \big(\langle x, x^*\rangle - \varphi(x)\big). \tag{1.1}$$

Moreover, the infimal convolution of φ and f is the function

$$\varphi \,\square\, f \colon \mathscr{X} \to [-\infty,+\infty] \colon x \mapsto \inf_{y\in\mathscr{X}} \big(\varphi(y) + f(x-y)\big). \tag{1.2}$$

Now suppose that $\mathscr{X}$ is a Hilbert space and set $q = (1/2)\|\cdot\|^2$. Then, for every $x \in \mathscr{X}$, there exists a unique point $p \in \mathscr{X}$ such that $(\varphi\,\square\, q)(x) = \varphi(p) + q(x-p)$; this point is denoted by $p = \operatorname{prox}_\varphi x$. The operator $\operatorname{prox}_\varphi \colon \mathscr{X} \to \mathscr{X}$ thus defined is called the proximity operator of φ.

Proposition 3 [23, 25] *Suppose that $\mathscr{X}$ is a Hilbert space, let $\varphi \in \Gamma_0(\mathscr{X})$, set $q = \|\cdot\|^2/2$, and let $x \in \mathscr{X}$. Then the following hold.*

(i) $q(x) = (\varphi\,\square\, q)(x) + (\varphi^*\,\square\, q)(x)$.
(ii) $x = \operatorname{prox}_\varphi x + \operatorname{prox}_{\varphi^*} x$.
(iii) $\langle \operatorname{prox}_\varphi x, \operatorname{prox}_{\varphi^*} x\rangle = \varphi\big(\operatorname{prox}_\varphi x\big) + \varphi^*\big(\operatorname{prox}_{\varphi^*} x\big)$.

Note that, if in Proposition 3 φ is the indicator function of a nonempty closed convex cone $K \subset \mathscr{X}$, i.e., $\varphi = \iota_K$ where

$$(\forall x \in \mathscr{X}) \quad \iota_K(x) = \begin{cases} 0, & \text{if } x \in K;\\ +\infty, & \text{if } x \notin K, \end{cases} \tag{1.3}$$

we recover Proposition 2.

The above hilbertian nonlinear decomposition principles have found many applications in optimization and in various other areas of applied mathematics (see for instance [9,12,13,14,15,16,17,18,21,28] and the references therein) and attempts have been made to extend them to more general Banach spaces. The main result in this direction is the following generalization of Proposition 2(ii)&(iii) in uniformly convex and uniformly smooth Banach spaces (see also [3,19,29,30] for alternate proofs and applications), where Π_C denotes the generalized projector onto a nonempty closed convex subset C of $\mathscr{X}$ [1], i.e., if J denotes the duality mapping of $\mathscr{X}$,

$$(\forall x \in \mathscr{X}) \quad \Pi_C x = \operatorname*{argmin}_{y\in C}\big(\|x\|^2 - 2\langle y, Jx\rangle + \|y\|^2\big). \tag{1.4}$$

Proposition 4 [2] *Suppose that $\mathscr{X}$ is uniformly convex and uniformly smooth, let $J\colon \mathscr{X} \to \mathscr{X}^*$ denote its duality mapping, which is characterized by*

$$(\forall x \in \mathscr{X}) \quad \|x\|^2 = \langle x, Jx\rangle = \|Jx\|_*^2, \tag{1.5}$$

let K be a nonempty closed convex cone in $\mathscr{X}$, and let $x \in \mathscr{X}$. Then the following hold.

(i) $x = P_K x + J^{-1}\big(\Pi_{K^\ominus}(Jx)\big)$.
(ii) $\langle P_K x, \Pi_{K^\ominus}(Jx)\rangle = 0$.

The objective of the present paper is to unify and extend the above results. To this end, we first discuss in Section 2 suitable notions of proximity in Banach spaces. Based on these, we propose our extension of Moreau's decomposition in Section 3. A feature of our analysis is to rely heavily on convex analytical tools, which allows us to derive our main result with simpler proofs than those utilized in the above special case.

2 Proximity in Banach spaces

Let $\varphi \in \Gamma_0(\mathscr{X})$. As seen in the Introduction, if $\mathscr{X}$ is a Hilbert space, Moreau's proximity operator is defined by

$$(\forall x \in \mathscr{X}) \quad \operatorname{prox}_\varphi x = \operatorname*{argmin}_{y\in\mathscr{X}}\Big(\varphi(y) + \frac{1}{2}\|x-y\|^2\Big). \tag{2.1}$$

In this section we discuss two extensions of this operator in Banach spaces. We recall that φ is coercive if $\lim_{\|y\|\to+\infty}\varphi(y) = +\infty$ and supercoercive if $\lim_{\|y\|\to+\infty}\varphi(y)/\|y\| = +\infty$. As usual, the subdifferential operator of φ is denoted by $\partial\varphi$. Finally, the strong relative interior of a convex set $C \subset \mathscr{X}$ is

$$\operatorname{sri} C = \left\{x \in C \;\middle|\; \bigcup_{\lambda>0}\lambda(C-x) = \overline{\operatorname{span}}(C-x)\right\}. \tag{2.2}$$

We shall also require the following facts.

Lemma 1 ([24,26]) *Let $f \in \Gamma_0(\mathscr{X})$ and let $x^* \in \mathscr{X}^*$. Then $f - x^*$ is coercive if and only if $x^* \in \operatorname{int}\operatorname{dom} f^*$.*

Lemma 2 ([5, Theorem 3.4]) *Let $f \in \Gamma_0(\mathscr{X})$ be supercoercive. Then $\operatorname{dom} f^* = \mathscr{X}^*$.*

Lemma 3 ([4]) *Let f and φ be functions in $\Gamma_0(\mathscr{X})$ such that $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom}\varphi)$. Then the following hold.*

(i) $(\varphi + f)^* = \varphi^* \,\square\, f^*$ *and the infimal convolution is exact everywhere:*

$$(\forall x^* \in \mathscr{X}^*)(\exists y^* \in \mathscr{X}^*) \quad (\varphi + f)^*(x^*) = \varphi^*(y^*) + f^*(x^* - y^*).$$

(ii) $\partial(\varphi + f) = \partial\varphi + \partial f$.

2.1 Legendre functions

We review the notion of a Legendre function, which was introduced in Euclidean spaces in [27] and extended to Banach spaces in [5] (see also [8] for further developments in the nonreflexive case).

Definition 1 [5, Definition 5.2] Let $f \in \Gamma_0(\mathscr{X})$. Then f is:

(i) essentially smooth, if ∂f is both locally bounded and single-valued on its domain;
(ii) essentially strictly convex, if $(\partial f)^{-1}$ is locally bounded on its domain and f is strictly convex on every convex subset of $\operatorname{dom}\partial f$;
(iii) a Legendre function, if it is both essentially smooth and essentially strictly convex.

Some key properties of Legendre functions are listed below.

Lemma 4 *Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function. Then the following hold.*

(i) *f^* is a Legendre function* [5, Corollary 5.5].
(ii) $\operatorname{dom}\partial f = \operatorname{int}\operatorname{dom} f \neq \varnothing$ *and f is Gâteaux differentiable on* $\operatorname{int}\operatorname{dom} f$ [5, Theorem 5.6].
(iii) $\nabla f\colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f^*$ *is bijective with inverse* $\nabla f^*\colon \operatorname{int}\operatorname{dom} f^* \to \operatorname{int}\operatorname{dom} f$ [5, Theorem 5.10].

2.2 D-proximity operators

In this subsection we discuss a notion of proximity based on Bregman distances investigated in [6] and which goes back to [10,31].

The first extension of (2.1) was investigated in [6]. Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function. The Bregman distance associated with f is

$$\begin{aligned} D_f\colon \mathscr{X} \times \mathscr{X} &\to [0,+\infty] \\ (y,x) &\mapsto \begin{cases} f(y) - f(x) - \langle y - x, \nabla f(x)\rangle, & \text{if } x \in \operatorname{int}\operatorname{dom} f; \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.3}$$

For every $\varphi \in \Gamma_0(\mathscr{X})$, we define the function $\varphi \diamond f\colon \mathscr{X} \to [-\infty,+\infty]$ by

$$(\forall x \in \mathscr{X}) \quad (\varphi \diamond f)(x) = \inf_{y\in\mathscr{X}} \big(\varphi(y) + D_f(y,x)\big). \tag{2.4}$$

The following proposition refines and complements some results of [6, Section 3.4].

Proposition 5 *Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function, let $\varphi \in \Gamma_0(\mathscr{X})$ be such that*

$$0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom}\varphi), \tag{2.5}$$

and let $x \in \operatorname{int}\operatorname{dom} f$. Suppose that one of the following holds.

(i) $\nabla f(x) \in \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)$.
(ii) $\operatorname{int}\operatorname{dom} f^* \subset \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)$.
(iii) *f is supercoercive.*
(iv) $\inf\varphi(\mathscr{X}) > -\infty$.

Then there exists a unique point $p \in \mathscr{X}$ such that $(\varphi \diamond f)(x) = \varphi(p) + D_f(p,x)$; moreover, p lies in $\operatorname{dom}\partial\varphi \cap \operatorname{int}\operatorname{dom} f$ and it is characterized by the inclusion

$$\nabla f(x) - \nabla f(p) \in \partial\varphi(p). \tag{2.6}$$

Proof Set $f_x\colon \mathscr{X} \to \,]-\infty,+\infty] : y \mapsto f(y) - \langle y, \nabla f(x)\rangle$. Then the minimizers of $\varphi + D_f(\cdot,x)$ coincide with those of $\varphi + f_x$ and our assumptions imply that

$$\varphi + f_x \in \Gamma_0(\mathscr{X}). \tag{2.7}$$

Now let $p \in \mathscr{X}$. It follows from (2.5), Lemma 3(ii), and Lemma 4(ii) that

$$\begin{aligned}
(\varphi \diamond f)(x) = \varphi(p) + D_f(p,x) &\Leftrightarrow p \text{ minimizes } \varphi + f_x \\
&\Leftrightarrow 0 \in \partial\big(\varphi + f_x\big)(p) \\
&\Leftrightarrow 0 \in \partial\varphi(p) + \partial f(p) - \nabla f(x) \\
&\Leftrightarrow 0 \in \partial\varphi(p) + \nabla f(p) - \nabla f(x) \\
&\Leftrightarrow \nabla f(x) - \nabla f(p) \in \partial\varphi(p) \qquad (2.8) \\
&\Rightarrow p \in \operatorname{dom}\partial\varphi \cap \operatorname{int}\operatorname{dom} f. \qquad (2.9)
\end{aligned}$$

Hence, the minimizers of $\varphi + f_x$ are in $\operatorname{int}\operatorname{dom} f$. However, since f is essentially strictly convex, it is strictly convex on $\operatorname{int}\operatorname{dom} f$ and so is therefore $\varphi + f_x$. This shows that $\varphi + f_x$ admits at most one minimizer. It remains to establish existence.

(i): It follows from (2.7) that, to show existence, it is enough to show that $\varphi + f_x$ is coercive [33, Theorem 2.5.1(ii)]. In view of Lemma 1, this is equivalent to showing that $\nabla f(x) \in \operatorname{int}\operatorname{dom}(f+\varphi)^*$. However, it follows from (2.5) and Lemma 3(i) that

$$\operatorname{int}\operatorname{dom}(f+\varphi)^* = \operatorname{int}\operatorname{dom}(f^* \,\square\, \varphi^*) = \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*). \tag{2.10}$$

(ii)⇒(i): Lemma 4(iii).

(iii)⇒(ii): By Lemma 2, $\operatorname{dom} f^* = \mathscr{X}^*$ and, since $\operatorname{dom}\varphi^* \neq \varnothing$, $\operatorname{int}\operatorname{dom} f^* \subset \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)$.

(iv)⇒(ii): We have $\inf\varphi(\mathscr{X}) > -\infty \Rightarrow \varphi^*(0) = -\inf\varphi(\mathscr{X}) < +\infty \Rightarrow 0 \in \operatorname{dom}\varphi^*$. Hence, $\operatorname{int}\operatorname{dom} f^* \subset \operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)$.

In view of Proposition 5 and Lemma 4(iii), the following is well defined.

Definition 2 Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function and let $\varphi \in \Gamma_0(\mathscr{X})$ be such that $0 \in \operatorname{sri}(\operatorname{dom} f - \operatorname{dom}\varphi)$. Set

$$E = (\operatorname{int}\operatorname{dom} f) \cap \big(\nabla f^* \big(\operatorname{int}(\operatorname{dom} f^* + \operatorname{dom}\varphi^*)\big)\big). \tag{2.11}$$

The D-proximity (or Bregman proximity) operator of φ relative to f is

$$\operatorname{bprox}_\varphi^f \colon E \to \operatorname{int}\operatorname{dom} f \colon x \mapsto \operatorname*{argmin}_{y\in\mathscr{X}} \big(\varphi(y) + D_f(y,x)\big). \tag{2.12}$$

Remark 1 In connection with Definition 2, let us make a couple of observations.

(i) It follows from Proposition 5 that, if $\operatorname{int}\operatorname{dom} f^* \subset \operatorname{int}(\operatorname{dom}\varphi^* + \operatorname{dom} f^*)$ (in particular if f is supercoercive or if $\inf\varphi(\mathscr{X}) > -\infty$), then $\operatorname{bprox}_\varphi^f \colon \operatorname{int}\operatorname{dom} f \to \operatorname{int}\operatorname{dom} f$.

(ii) Suppose that $\mathscr{X}$ is hilbertian and that $f = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathscr{X})$. Then $\varphi \diamond f = \varphi \,\square\, f$ and $\operatorname{bprox}_\varphi^f = \operatorname{prox}_\varphi$.

2.3 Anisotropic proximity operators

An alternative extension of the notion of proximity can be obtained by replacing the function $\|\cdot\|^2/2$ in (2.1) by a Legendre function f. This type of construction goes back to [20].

Proposition 6 *Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function, let $\varphi \in \Gamma_0(\mathscr{X})$ be such that*

$$0 \in \operatorname{sri}(\operatorname{dom} f^* - \operatorname{dom}\varphi^*), \tag{2.13}$$

and let $x \in \operatorname{sri}(\operatorname{dom} f + \operatorname{dom}\varphi)$. Then there exists a unique point $p \in \mathscr{X}$ such that $(\varphi \,\square\, f)(x) = \varphi(p) + f(x-p)$; moreover, p is characterized by the inclusion

$$\nabla f(x-p) \in \partial\varphi(p). \tag{2.14}$$

Proof Using (2.13) and Lemma 3(i), we obtain

$$(\varphi^* + f^*)^* = \varphi^{**} \,\square\, f^{**} = \varphi \,\square\, f \tag{2.15}$$

and the fact that the infimum in the infimal convolution is attained everywhere. On the other hand, since $x \in \operatorname{sri}(\operatorname{dom} f + \operatorname{dom}\varphi)$, we have

$$0 \in \operatorname{sri}\big(\operatorname{dom}\varphi - (x - \operatorname{dom} f)\big) = \operatorname{sri}\big(\operatorname{dom}\varphi - \operatorname{dom} f(x-\cdot)\big). \tag{2.16}$$

Consequently, by Lemma 3(ii),

$$\partial\big(\varphi + f(x-\cdot)\big) = \partial\varphi + \partial f(x-\cdot). \tag{2.17}$$

Now let $p \in \mathscr{X}$. It follows from (2.17) and Lemma 4(ii) that

$$\begin{aligned} p \text{ minimizes } \varphi + f(x-\cdot) &\Leftrightarrow 0 \in \partial\big(\varphi + f(x-\cdot)\big)(p) \\ &\Leftrightarrow 0 \in \partial\varphi(p) - \partial f(x-p) \\ &\Leftrightarrow 0 \in \partial\varphi(p) - \nabla f(x-p) \\ &\Leftrightarrow \nabla f(x-p) \in \partial\varphi(p) \qquad (2.18) \\ &\Rightarrow x-p \in \operatorname{int}\operatorname{dom} f. \qquad (2.19) \end{aligned}$$

To show uniqueness, suppose that p and q are two distinct minimizers of $\varphi + f(x-\cdot)$. Then $(\varphi \,\square\, f)(x) = \varphi(p) + f(x-p) = \varphi(q) + f(x-q)$ and, by (2.19), $x-p$ and $x-q$ lie in $\operatorname{int}\operatorname{dom} f$. Now let $r = (1/2)(p+q)$ and suppose that $p \neq q$. Lemma 4(ii) asserts that f is strictly convex on the convex set $\operatorname{int}\operatorname{dom} f = \operatorname{dom}\partial f$. Therefore, invoking the convexity of φ,

$$\begin{aligned} (\varphi \,\square\, f)(x) &\leq \varphi(r) + f(x-r) \\ &< \frac{1}{2}\big(\varphi(p) + \varphi(q)\big) + \frac{1}{2}\big(f(x-p) + f(x-q)\big) \\ &= (\varphi \,\square\, f)(x), \qquad (2.20) \end{aligned}$$

which is impossible.

Using Proposition 6, we can now introduce the anisotropic proximity operator of φ.

Definition 3 Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function and let $\varphi \in \Gamma_0(\mathscr{X})$ be such that $0 \in \operatorname{sri}(\operatorname{dom} f^* - \operatorname{dom}\varphi^*)$. Set

$$E = \operatorname{sri}(\operatorname{dom} f + \operatorname{dom}\varphi). \qquad (2.21)$$

The anisotropic proximity operator of φ relative to f is

$$\operatorname{aprox}_{\varphi}^{f} \colon E \to \mathscr{X} \colon x \mapsto \operatorname*{argmin}_{y \in \mathscr{X}} \big(\varphi(y) + f(x-y)\big). \qquad (2.22)$$

Remark 2 Suppose that $\mathscr{X}$ is hilbertian and that $f = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathscr{X})$. Then $\operatorname{aprox}_{\varphi}^{f} = \operatorname{prox}_{\varphi}$.

3 Main result

In the previous section we have described two extensions of the classical proximity operator. Our main result is a generalization of Moreau's decomposition (Proposition 3) in Banach spaces which involves a mix of these two extensions.

Theorem 1 *Let $f \in \Gamma_0(\mathscr{X})$ be a Legendre function, let $\varphi \in \Gamma_0(\mathscr{X})$ be such that*

$$0 \in \operatorname{sri}(\operatorname{dom} f^* - \operatorname{dom}\varphi^*), \qquad (3.1)$$

and let $x \in (\operatorname{int}\operatorname{dom} f) \cap \operatorname{int}(\operatorname{dom} f + \operatorname{dom}\varphi)$. Then the following hold.

(i) $f(x)=(\varphi\,\square\, f)(x)+(\varphi^*\diamond f^*)\big(\nabla f(x)\big)$.
(ii) $x=\operatorname{aprox}_{\varphi}^{f}x+\nabla f^*\big(\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big)\big)$.
(iii) $\big\langle \operatorname{aprox}_{\varphi}^{f}x,\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big)\big\rangle=\varphi\big(\operatorname{aprox}_{\varphi}^{f}x\big)+\varphi^*\big(\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big)\big)$.
(iv) $\big\langle \operatorname{aprox}_{\varphi}^{f}x,\nabla f\big(x-\operatorname{aprox}_{\varphi}^{f}x\big)\big\rangle=\varphi\big(\operatorname{aprox}_{\varphi}^{f}x\big)+\varphi^*\big(\nabla f\big(x-\operatorname{aprox}_{\varphi}^{f}x\big)\big)$.

Proof Since $x\in\operatorname{int}(\operatorname{dom} f+\operatorname{dom}\varphi)$, Lemma 4(iii) yields

$$x\in\operatorname{sri}(\operatorname{dom} f+\operatorname{dom}\varphi)\quad\text{and}\quad\nabla f^*\big(\nabla f(x)\big)\in\operatorname{int}\big(\operatorname{dom} f^{**}+\operatorname{dom}\varphi^{**}\big). \tag{3.2}$$

Hence, it follows from Proposition 6 that $\operatorname{aprox}_{\varphi}^{f}x$ is well defined and, from Lemma 4(i) and Proposition 5(i) (applied to f^* and φ^*), that $\nabla f^*(\operatorname{bprox}_{\varphi^*}^{f^*}(\nabla f(x)))$ is well defined. In addition,

$$(\varphi\,\square\, f)(x)\in\mathbb{R}\quad\text{and}\quad(\varphi^*\diamond f^*)\big(\nabla f(x)\big)\in\mathbb{R}. \tag{3.3}$$

(i): It follows from (2.3), Lemma 4(iii), and the Fenchel-Young identity [33, Theorem 2.4.2(iii)] that

$$\begin{aligned}(\forall x^*\in\mathscr{X}^*)\quad D_{f^*}\big(x^*,\nabla f(x)\big)&=f^*(x^*)-f^*\big(\nabla f(x)\big)-\langle x^*-\nabla f(x),x\rangle_*\\&=f^*(x^*)+f(x)-\langle x^*,x\rangle_*.\end{aligned} \tag{3.4}$$

This, (2.4), (3.1), and Lemma 3(i) imply that

$$\begin{aligned}(\varphi^*\diamond f^*)\big(\nabla f(x)\big)&=\inf_{x^*\in\mathscr{X}^*}\big(\varphi^*(x^*)+f^*(x^*)+f(x)-\langle x^*,x\rangle_*\big)\\&=f(x)-\sup_{x^*\in\mathscr{X}^*}\big(\langle x^*,x\rangle_*-\varphi^*(x^*)-f^*(x^*)\big)\\&=f(x)-(\varphi^*+f^*)^*(x)\\&=f(x)-(\varphi\,\square\, f)(x).\end{aligned} \tag{3.5}$$

In view of (3.3), we obtain the announced identity.

(ii): Let $p\in\mathscr{X}$. Using Proposition 6, Lemma 4(iii), and Proposition 5(i), we obtain

$$\begin{aligned}p=\operatorname{aprox}_{\varphi}^{f}x&\Leftrightarrow\nabla f(x-p)\in\partial\varphi(p) &(3.6)\\&\Leftrightarrow p\in\partial\varphi^*\big(\nabla f(x-p)\big)\\&\Leftrightarrow\nabla f^*\big(\nabla f(x)\big)-\nabla f^*\big(\nabla f(x-p)\big)\in\partial\varphi^*\big(\nabla f(x-p)\big)\\&\Leftrightarrow\nabla f(x-p)=\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big) &(3.7)\\&\Leftrightarrow x-p=\nabla f^*\big(\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big)\big). &(3.8)\end{aligned}$$

(iii): Set $p=\operatorname{aprox}_{\varphi}^{f}x$. As seen in (3.7) and (3.6),

$$\operatorname{bprox}_{\varphi^*}^{f^*}\big(\nabla f(x)\big)=\nabla f(x-p)\in\partial\varphi(p). \tag{3.9}$$

Hence, the Fenchel-Young identity yields

$$\begin{aligned}\langle p, \operatorname{bprox}^{f^*}_{\varphi^*}(\nabla f(x))\rangle &= \langle p, \nabla f(x-p)\rangle \\ &= \varphi(p) + \varphi^*\big(\nabla f(x-p)\big) \\ &= \varphi(p) + \varphi^*\big(\operatorname{bprox}^{f^*}_{\varphi^*}(\nabla f(x))\big).\end{aligned} \tag{3.10}$$

(iv): This follows at once from (iii) and (3.9).

Remark 3 An instance of Theorem 1(iv) in which f and f^* are real-valued appears in [32, Proposition 1].

Theorem 1 provides a range of new decomposition schemes, even in the case when $\mathscr{X}$ is a Hilbert space. Thus, in the following result, we obtain a new hilbertian frame decomposition principle (for background on frames and their applications, see [11]).

Corollary 1 *Suppose that $\mathscr{X}$ is a separable Hilbert space, let I be a countable set, and let $(e_i)_{i\in I}$ be a frame in $\mathscr{X}$, i.e.,*

$$(\exists\, \alpha \in]0,+\infty[)(\exists\, \beta \in]0,+\infty[)(\forall x \in \mathscr{X}) \quad \alpha\|x\|^2 \leq \sum_{i\in I} |\langle x, e_i\rangle|^2 \leq \beta\|x\|^2. \tag{3.11}$$

Let $S\colon \mathscr{X} \to \mathscr{X} : x \mapsto \sum_{i\in I}\langle x, e_i\rangle e_i$ be the associated frame operator and let $(e_i^*)_{i\in I} = (S^{-1}e_i)_{i\in I}$ be the associated canonical dual frame. Furthermore, let $\varphi \in \Gamma_0(\mathscr{X})$, let $x \in \mathscr{X}$, and set

$$a(x) = \operatorname*{argmin}_{y\in\mathscr{X}} \left(\varphi(y) + \frac{1}{2}\sum_{i\in I} |\langle x-y, e_i\rangle|^2\right) \tag{3.12}$$

and

$$b(x) = \operatorname*{argmin}_{x^*\in\mathscr{X}} \left(\varphi^*(x^*) - \langle x^*, x\rangle + \frac{1}{2}\sum_{i\in I} |\langle x^*, e_i^*\rangle|^2\right). \tag{3.13}$$

Then $x = a(x) + \sum_{i\in I}\langle b(x), e_i^*\rangle e_i^*$.

Proof Set $f\colon \mathscr{X} \to \mathbb{R}\colon x \mapsto (1/2)\sum_{i\in I}|\langle x, e_i\rangle|^2$. It is easily seen that f is Fréchet differentiable on $\mathscr{X}$ with $\nabla f = S$. It therefore follows from [5, Theorem 5.6] that f is essentially smooth. Now fix $x^* \in \mathscr{X}$. Since the frame operator of $(e_i^*)_{i\in I}$ is S^{-1} [11, Lemma 5.1.6], we have

$$\langle S^{-1}x^*, x^*\rangle = \Big\langle \sum_{i\in I}\langle x^*, e_i^*\rangle e_i^*, x^*\Big\rangle = \sum_{i\in I}|\langle x^*, e_i^*\rangle|^2 = 2f(S^{-1}x^*). \tag{3.14}$$

Now set $g\colon \mathscr{X} \to \mathbb{R}\colon x \mapsto f(x) - \langle x, x^*\rangle$. Then g is a differentiable convex function and $\nabla g\colon x \mapsto Sx - x^*$ vanishes at $x = S^{-1}x^*$. Hence, using (3.14), we obtain

$$f^*(x^*) = -\min_{x\in\mathscr{X}} g(x) = \langle S^{-1}x^*, x^*\rangle - f(S^{-1}x^*) = f(S^{-1}x^*) = \frac{1}{2}\sum_{i\in I}|\langle x^*, e_i^*\rangle|^2. \tag{3.15}$$

Hence, as above, f^* is Fréchet differentiable on $\mathscr{X}$ with $\nabla f^* = S^{-1}$ and, in turn, essentially smooth, which makes f essentially strictly convex [5, Theorem 5.4]. Altogether, f is a Legendre function with

$$\operatorname{dom} f = \mathscr{X}, \quad \operatorname{dom} f^* = \mathscr{X}, \quad \nabla f = S, \quad \text{and} \quad \nabla f^* = S^{-1}. \tag{3.16}$$

Moreover, it follows from (2.12), (2.22), (3.16), Lemma 4(iii), (3.12), (3.13), and (3.15) that

$$\operatorname{bprox}_{\varphi^*}^{f^*}(\nabla f(x)) = b(x) \quad \text{and} \quad \operatorname{aprox}_{\varphi}^{f}(x) = a(x). \tag{3.17}$$

The result is therefore an application of Theorem 1(ii).

Remark 4 Corollary 1 can be regarded as an extension of Moreau's decomposition principle in separable Hilbert spaces. Indeed, in the special case when $(e_i)_{i\in I}$ is an orthonormal basis in Corollary 1, we recover Proposition 3(ii).

The next application is set in uniformly convex and uniformly smooth Banach spaces.

Corollary 2 *Suppose that $\mathscr{X}$ is uniformly convex and uniformly smooth, let J be its duality mapping, set $q = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathscr{X})$. Then $q^* = \|\cdot\|_*^2/2$ and the following hold for every $x \in \mathscr{X}$.*

(i) $q(x) = (\varphi \,\square\, q)(x) + (\varphi^* \,\diamond\, q^*)(Jx)$.
(ii) $x = \operatorname{aprox}_{\varphi}^{q} x + J^{-1}\big(\operatorname{bprox}_{\varphi^*}^{q^*}(Jx)\big)$.
(iii) $\big\langle \operatorname{aprox}_{\varphi}^{q} x, \operatorname{bprox}_{\varphi^*}^{q^*}(Jx)\big\rangle = \varphi\big(\operatorname{aprox}_{\varphi}^{q} x\big) + \varphi^*\big(\operatorname{bprox}_{\varphi^*}^{q^*}(Jx)\big)$.
(iv) $\big\langle \operatorname{aprox}_{\varphi}^{q} x, J\big(x - \operatorname{aprox}_{\varphi}^{q} x\big)\big\rangle = \varphi\big(\operatorname{aprox}_{\varphi}^{q} x\big) + \varphi^*\big(J\big(x - \operatorname{aprox}_{\varphi}^{q} x\big)\big)$.

Proof This is an application of Theorem 1 with $f = q$. Indeed, $\operatorname{dom} f = \mathscr{X}$, $\operatorname{dom} f^* = \mathscr{X}^*$, and $\nabla f = J$.

In particular, if $\mathscr{X}$ is a Hilbert space in Corollary 2, if follows from Remark 1(ii) and Remark 2 that we recover Moreau's decomposition principle (Proposition 3) and a fortiori Propositions 1 and 2. Another noteworthy instance of Corollary 2 is when $\varphi = \iota_K$, where K is a nonempty closed convex cone in $\mathscr{X}$. In this case, $\varphi^* = \iota_{K^\ominus}$, $\operatorname{aprox}_{\varphi}^{q} = P_K$, and we derive from (1.4) and (1.5) that $\operatorname{bprox}_{\varphi}^{q} = \Pi_K$. Hence, Corollary 2(ii)&(iii) yields Proposition 4.

Remark 5 Consider the setting of Theorem 1 and set $A = \partial\varphi$. Then, by Rockafellar's theorem, A is a maximally monotone operator [33, Theorem 3.1.11]. Moreover, it follows from (2.14), Lemma 4(iii), and (2.6) that we can rewrite Theorem 1(ii) as

$$x = (\mathrm{Id} + \nabla f^* \circ A)^{-1} x + \nabla f^* \circ \big(\nabla f^* + A^{-1}\big)^{-1} x, \tag{3.18}$$

where Id is the identity operator on $\mathscr{X}$. The results of [6, Section 3.3] suggest that this decomposition holds for more general maximally monotone operators $A\colon \mathscr{X} \to 2^{\mathscr{X}^*}$. If $\mathscr{X}$ is a Hilbert space and $f = \|\cdot\|^2/2$, (3.18) yields the well-known resolvent identity $\mathrm{Id} = (\mathrm{Id} + A)^{-1} + (\mathrm{Id} + A^{-1})^{-1}$, which is true for any maximally monotone operator A [7, Proposition 23.18].

Acknowledgements The authors thank one of the referees for making some valuable comments.

References

1. Ya. I. Alber, Metric and generalized projection operators in Banach spaces: properties and applications, In: *Theory and Applications of Nonlinear Operators of Accretive and Monotone Type,* (A. Kartsatos, Ed.), 15–50. Marcel Dekker, New York (1996)
2. Ya. I. Alber, Decomposition theorems in Banach spaces, in: *Operator Theory and Its Applications*, 77–93. Fields Institute Communications **25**, AMS, Providence, RI (2000)
3. Ya. I. Alber, James orthogonality and orthogonal decompositions of Banach spaces, *J. Math. Anal. Appl.* **312**, 330–342 (2005)
4. H. Attouch and H. Brézis, Duality for the sum of convex functions in general Banach spaces, in: *Aspects of Mathematics and Its Applications*, 125–133. North-Holland, Amsterdam (1986)
5. H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, *Commun. Contemp. Math.* **3**, 615–647 (2001)
6. H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM J. Control Optim.*, **42**, 596–636 (2003)
7. H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, New York (2011)
8. J. M. Borwein and J. D. Vanderwerff, Convex functions of Legendre type in general Banach spaces, *J. Convex Anal.* **8**, 569–581 (2001)
9. B. Brogliato, A. Daniilidis, C. Lemaréchal, and V. Acary, On the equivalence between complementarity systems, projected systems and differential inclusions, *Systems Control Lett.* **55**, 45–51 (2006)
10. Y. Censor and S. A. Zenios, Proximal minimization algorithm with D-functions, *J. Optim. Theory Appl.*, **73**, 451–464 (1992)
11. O. Christensen, *Frames and Bases – An Introductory Course.* Birkhäuser, Boston, MA (2008)
12. W. D. Collins, Dual extremum principles and Hilbert space decompositions, in: *Duality and Complementarity in Mechanics of Solids*, 351–418. Ossolineum, Wrocław (1979)
13. P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued Var. Anal.* **18**, 373–404 (2010)
14. P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.* **18**, 1351–1376 (2007)
15. P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.* **4**, 1168–1200 (2005)
16. S.-P. Han and O. L. Mangasarian, Conjugate cone characterization of positive definite and semidefinite matrices, *Lin. Algebra Appl.* **56**, 89–103 (1984)
17. J.-B. Hiriart-Urruty and Ph. Plazanet, Moreau's decomposition theorem revisited, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**, 325–338 (1989)
18. J.-B. Hiriart-Urruty and A. Seeger, A variational approach to copositive matrices, *SIAM Rev.* **52**, 593–629 (2010)
19. Y. H. Hu and W. Song, Weak sharp solutions for variational inequalities in Banach spaces, *J. Math. Anal. Appl.* **374**, 118–132 (2011)
20. C. Lescarret, Applications "prox" dans un espace de Banach, *C. R. Acad. Sci. Paris Sér. A Math.*, **265**, 676–678 (1967)
21. Y. Lucet, What shape is your conjugate? A survey of computational convex analysis and its applications, *SIAM Rev.* **52**, 505–542 (2010)
22. J.-J. Moreau, Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires, *C. R. Acad. Sci. Paris Sér. A Math.*, **255**, 238–240 (1962)
23. J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.*, **255**, 2897–2899 (1962)
24. J.-J. Moreau, Sur la fonction polaire d'une fonction semi-continue supérieurement, *C. R. Acad. Sci. Paris Sér. A Math.*, **258**, 1128–1130 (1964)
25. J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France* **93**, 273–299 (1965)
26. R. T. Rockafellar, Level sets and continuity of conjugate convex functions, *Trans. Amer. Math. Soc.*, **123**, 46–63 (1966)
27. R. T. Rockafellar, *Convex Analysis.* Princeton University Press, Princeton, New Jersey (1970)
28. R. T. Rockafellar, Moreau's proximal mappings and convexity in Hamilton-Jacobi theory, in: *Nonsmooth Mechanics and Analysis*, 3–12. Springer, New York (2006)

29. F. Schöpfer, T. Schuster, and A. K. Louis, Metric and Bregman projections onto affine subspaces and their computation via sequential subspace optimization methods, *J. Inverse Ill-Posed Probl.* **16**, 479–506 (2008)
30. W. Song and Z. Cao, The generalized decomposition theorem in Banach spaces and its applications, *J. Approx. Theory* **129**, 167–181 (2004)
31. M. Teboulle, Entropic proximal mappings with applications to nonlinear programming, *Math. Oper. Res.* **17**, 670–690 (1992)
32. D. Wexler, Prox-mappings associated with a pair of Legendre conjugate functions, *Rev. Française Automat. Informat. Recherche Opérationnelle* **7**, 39–65 (1973)
33. C. Zălinescu, *Convex Analysis in General Vector Spaces*. World Scientific, River Edge, NJ (2002)