



Moreau's decomposition in Banach spaces

Patrick Louis Combettes, Noli Reyes

► To cite this version:

Patrick Louis Combettes, Noli Reyes. Moreau's decomposition in Banach spaces. Mathematical Programming, 2013, 139, pp.103 - 114. 10.1007/s10107-013-0663-y . hal-01076974

HAL Id: hal-01076974

<https://hal.science/hal-01076974>

Submitted on 23 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Moreau's Decomposition in Banach Spaces

Patrick L. Combettes¹ and Noli N. Reyes²

¹UPMC Université Paris 06
Laboratoire Jacques-Louis Lions – UMR CNRS 7598
75005 Paris, France
plc@math.jussieu.fr

²University of the Philippines – Diliman
Institute of Mathematics
Quezon City, 1101 Philippines
noli@math.upd.edu.ph

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.

Key words. Banach space, Bregman distance, convex optimization, infimal convolution, Legendre function, Moreau's decomposition, proximity operator.

1 Introduction

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

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

Proposition 1.1 *Suppose that \mathcal{X} is a Hilbert space, let V be a closed vector subspace of \mathcal{X} with orthogonal complement V^\perp , and let $x \in \mathcal{X}$. 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 1.2 [22] *Suppose that \mathcal{X} is a Hilbert space, let K be a nonempty closed convex cone in \mathcal{X} , and let $x \in \mathcal{X}$. 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 $\Gamma_0(\mathcal{X})$. The conjugate of φ is the function φ^* in $\Gamma_0(\mathcal{X}^*)$ defined by

$$\varphi^*: \mathcal{X}^* \rightarrow]-\infty, +\infty] : x^* \mapsto \sup_{x \in \mathcal{X}} (\langle x, x^* \rangle - \varphi(x)). \quad (1.1)$$

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

$$\varphi \square f: \mathcal{X} \rightarrow [-\infty, +\infty] : x \mapsto \inf_{y \in \mathcal{X}} (\varphi(y) + f(x - y)). \quad (1.2)$$

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

Proposition 1.3 [23, 25] *Suppose that \mathcal{X} is a Hilbert space, let $\varphi \in \Gamma_0(\mathcal{X})$, set $q = \|\cdot\|^2/2$, and let $x \in \mathcal{X}$. Then the following hold.*

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

Note that, if in Proposition 1.3 φ is the indicator function of a nonempty closed convex cone $K \subset \mathcal{X}$, i.e., $\varphi = \iota_K$ where

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

we recover Proposition 1.2.

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 1.2(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 \mathcal{X} [1], i.e., if J denotes the duality mapping of \mathcal{X} ,

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

Proposition 1.4 [2] *Suppose that \mathcal{X} is uniformly convex and uniformly smooth, let $J: \mathcal{X} \rightarrow \mathcal{X}^*$ denote its duality mapping, which is characterized by*

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

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

- (i) $x = P_K x + J^{-1}(\Pi_{K^\ominus}(Jx))$.
- (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 $\varphi \in \Gamma_0(\mathcal{X})$. As seen in the Introduction, if \mathcal{X} is a Hilbert space, Moreau's proximity operator is defined by

$$(\forall x \in \mathcal{X}) \quad \text{prox}_\varphi x = \underset{y \in \mathcal{X}}{\operatorname{argmin}} \left(\varphi(y) + \frac{1}{2} \|x - y\|^2 \right). \quad (2.1)$$

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

$$\operatorname{sri} C = \left\{ x \in C \mid \bigcup_{\lambda > 0} \lambda(C - x) = \overline{\operatorname{span}}(C - x) \right\}. \quad (2.2)$$

We shall also require the following facts.

Lemma 2.1 ([24, 26]) *Let $f \in \Gamma_0(\mathcal{X})$ and let $x^* \in \mathcal{X}^*$. Then $f - x^*$ is coercive if and only if $x^* \in \operatorname{int} \operatorname{dom} f^*$.*

Lemma 2.2 ([5, Theorem 3.4]) *Let $f \in \Gamma_0(\mathcal{X})$ be supercoercive. Then $\operatorname{dom} f^* = \mathcal{X}^*$.*

Lemma 2.3 ([4]) *Let f and φ be functions in $\Gamma_0(\mathcal{X})$ 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: $(\forall x^* \in \mathcal{X}^*)(\exists y^* \in \mathcal{X}^*)$
 $(\varphi + f)^*(x^*) = \varphi^*(y^*) + f^*(x^* - y^*)$.
- (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 2.4 [5, Definition 5.2] *Let $f \in \Gamma_0(\mathcal{X})$. 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 2.5 *Let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function. Then the following hold.*

- (i) f^* is a Legendre function [5, Corollary 5.5].
- (ii) $\text{dom } \partial f = \text{int dom } f \neq \emptyset$ and f is Gâteaux differentiable on $\text{int dom } f$ [5, Theorem 5.6].
- (iii) $\nabla f: \text{int dom } f \rightarrow \text{int dom } f^*$ is bijective with inverse $\nabla f^*: \text{int dom } f^* \rightarrow \text{int dom } f$ [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 $f \in \Gamma_0(\mathcal{X})$ be a Legendre function. The Bregman distance associated with f is

$$D_f: \mathcal{X} \times \mathcal{X} \rightarrow [0, +\infty]$$

$$(y, x) \mapsto \begin{cases} f(y) - f(x) - \langle y - x, \nabla f(x) \rangle, & \text{if } x \in \text{int dom } f; \\ +\infty, & \text{otherwise.} \end{cases} \quad (2.3)$$

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

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

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

Proposition 2.6 *Let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $\varphi \in \Gamma_0(\mathcal{X})$ be such that*

$$0 \in \text{sri}(\text{dom } f - \text{dom } \varphi), \quad (2.5)$$

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

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

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

$$\nabla f(x) - \nabla f(p) \in \partial \varphi(p). \quad (2.6)$$

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

$$\varphi + f_x \in \Gamma_0(\mathcal{X}). \quad (2.7)$$

Now let $p \in \mathcal{X}$. It follows from (2.5), Lemma 2.3(ii), and Lemma 2.5(ii) that

$$\begin{aligned} (\varphi \diamond f)(x) = \varphi(p) + D_f(p, x) &\Leftrightarrow p \text{ minimizes } \varphi + f_x \\ &\Leftrightarrow 0 \in \partial(\varphi + f_x)(p) \\ &\Leftrightarrow 0 \in \partial\varphi(p) + \partial f(p) - \nabla f(x) \\ &\Leftrightarrow 0 \in \partial\varphi(p) + \nabla f(p) - \nabla f(x) \\ &\Leftrightarrow \nabla f(x) - \nabla f(p) \in \partial\varphi(p) \end{aligned} \quad (2.8)$$

$$\Rightarrow p \in \text{dom } \partial\varphi \cap \text{int dom } f. \quad (2.9)$$

Hence, the minimizers of $\varphi + f_x$ are in $\text{int dom } f$. However, since f is essentially strictly convex, it is strictly convex on $\text{int dom } f$ 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 2.1, this is equivalent to showing that $\nabla f(x) \in \text{int dom } (f + \varphi)^*$. However, it follows from (2.5) and Lemma 2.3(i) that

$$\text{int dom } (f + \varphi)^* = \text{int dom } (f^* \square \varphi^*) = \text{int}(\text{dom } f^* + \text{dom } \varphi^*). \quad (2.10)$$

(ii) \Rightarrow (i): Lemma 2.5(iii).

(iii) \Rightarrow (ii): By Lemma 2.2, $\text{dom } f^* = \mathcal{X}^*$ and, since $\text{dom } \varphi^* \neq \emptyset$, $\text{int dom } f^* \subset \text{int}(\text{dom } f^* + \text{dom } \varphi^*)$.

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

In view of Proposition 2.6 and Lemma 2.5(iii), the following is well defined.

Definition 2.7 Let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function and let $\varphi \in \Gamma_0(\mathcal{X})$ be such that $0 \in \text{sri}(\text{dom } f - \text{dom } \varphi)$. Set

$$E = (\text{int dom } f) \cap (\nabla f^*(\text{int}(\text{dom } f^* + \text{dom } \varphi^*))). \quad (2.11)$$

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

$$\text{bprox}_\varphi^f: E \rightarrow \text{int dom } f: x \mapsto \underset{y \in \mathcal{X}}{\text{argmin}} (\varphi(y) + D_f(y, x)). \quad (2.12)$$

Remark 2.8 In connection with Definition 2.7, let us make a couple of observations.

- (i) It follows from Proposition 2.6 that, if $\text{int dom } f^* \subset \text{int}(\text{dom } \varphi^* + \text{dom } f^*)$ (in particular if f is supercoercive or if $\inf \varphi(\mathcal{X}) > -\infty$), then $\text{bprox}_\varphi^f: \text{int dom } f \rightarrow \text{int dom } f$.
- (ii) Suppose that \mathcal{X} is hilbertian and that $f = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathcal{X})$. Then $\varphi \diamond f = \varphi \square f$ and $\text{bprox}_\varphi^f = \text{prox}_\varphi$.

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

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

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

$$\nabla f(x - p) \in \partial \varphi(p). \quad (2.14)$$

Proof. Using (2.13) and Lemma 2.3(i), we obtain

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

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

$$0 \in \text{sri}(\text{dom } \varphi - (x - \text{dom } f)) = \text{sri}(\text{dom } \varphi - \text{dom } f(x - \cdot)). \quad (2.16)$$

Consequently, by Lemma 2.3(ii),

$$\partial(\varphi + f(x - \cdot)) = \partial \varphi + \partial f(x - \cdot). \quad (2.17)$$

Now let $p \in \mathcal{X}$. It follows from (2.17) and Lemma 2.5(ii) that

$$\begin{aligned} p \text{ minimizes } \varphi + f(x - \cdot) &\Leftrightarrow 0 \in \partial(\varphi + f(x - \cdot))(p) \\ &\Leftrightarrow 0 \in \partial \varphi(p) - \partial f(x - p) \\ &\Leftrightarrow 0 \in \partial \varphi(p) - \nabla f(x - p) \\ &\Leftrightarrow \nabla f(x - p) \in \partial \varphi(p) \end{aligned} \quad (2.18)$$

$$\Rightarrow x - p \in \text{int dom } f. \quad (2.19)$$

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 $\text{int dom } f$. Now let $r = (1/2)(p + q)$ and suppose that $p \neq q$. Lemma 2.5(ii) asserts that f is strictly convex on the convex set $\text{int dom } f = \text{dom } \partial f$. Therefore, invoking the convexity of φ ,

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

which is impossible. \square

Using Proposition 2.9, we can now introduce the anisotropic proximity operator of φ .

Definition 2.10 Let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function and let $\varphi \in \Gamma_0(\mathcal{X})$ be such that $0 \in \text{sri}(\text{dom } f^* - \text{dom } \varphi^*)$. Set

$$E = \text{sri}(\text{dom } f + \text{dom } \varphi). \quad (2.21)$$

The anisotropic proximity operator of φ relative to f is

$$\text{aprox}_\varphi^f: E \rightarrow \mathcal{X}: x \mapsto \underset{y \in \mathcal{X}}{\text{argmin}} (\varphi(y) + f(x - y)). \quad (2.22)$$

Remark 2.11 Suppose that \mathcal{X} is hilbertian and that $f = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathcal{X})$. Then $\text{aprox}_\varphi^f = \text{prox}_\varphi$.

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 1.3) in Banach spaces which involves a mix of these two extensions.

Theorem 3.1 Let $f \in \Gamma_0(\mathcal{X})$ be a Legendre function, let $\varphi \in \Gamma_0(\mathcal{X})$ be such that

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

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

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

Proof. Since $x \in \text{int}(\text{dom } f + \text{dom } \varphi)$, Lemma 2.5(iii) yields

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

Hence, it follows from Proposition 2.9 that $\text{aprox}_\varphi^f x$ is well defined and, from Lemma 2.5(i) and Proposition 2.6(i) (applied to f^* and φ^*), that $\nabla f^*(\text{bprox}_{\varphi^*}^{f^*}(\nabla f(x)))$ is well defined. In addition,

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

(i): It follows from (2.3), Lemma 2.5(iii), and the Fenchel-Young identity [33, Theorem 2.4.2(iii)] that

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

This, (2.4), (3.1), and Lemma 2.3(i) imply that

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

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

(ii): Let $p \in \mathcal{X}$. Using Proposition 2.9, Lemma 2.5(iii), and Proposition 2.6(i), we obtain

$$p = \text{aprox}_{\varphi}^f x \Leftrightarrow \nabla f(x - p) \in \partial\varphi(p) \tag{3.6}$$

$$\begin{aligned}
&\Leftrightarrow p \in \partial\varphi^*(\nabla f(x - p)) \\
&\Leftrightarrow \nabla f^*(\nabla f(x)) - \nabla f^*(\nabla f(x - p)) \in \partial\varphi^*(\nabla f(x - p)) \\
&\Leftrightarrow \nabla f(x - p) = \text{bprox}_{\varphi^*}^{f^*}(\nabla f(x))
\end{aligned} \tag{3.7}$$

$$\Leftrightarrow x - p = \nabla f^*(\text{bprox}_{\varphi^*}^{f^*}(\nabla f(x))). \tag{3.8}$$

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

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

Hence, the Fenchel-Young identity yields

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

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

Remark 3.2 An instance of Theorem 3.1(iv) in which f and f^* are real-valued appears in [32, Proposition 1].

Theorem 3.1 provides a range of new decomposition schemes, even in the case when \mathcal{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 3.3 Suppose that \mathcal{X} is a separable Hilbert space, let I be a countable set, and let $(e_i)_{i \in I}$ be a frame in \mathcal{X} , i.e.,

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

Let $S: \mathcal{X} \rightarrow \mathcal{X}: x \mapsto \sum_{i \in I} \langle x, e_i \rangle e_i$ 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 $\varphi \in \Gamma_0(\mathcal{X})$, let $x \in \mathcal{X}$, and set

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

and

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

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

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

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

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

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

Hence, as above, f^* is Fréchet differentiable on \mathcal{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

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

Moreover, it follows from (2.12), (2.22), (3.16), Lemma 2.5(iii), (3.12), (3.13), and (3.15) that

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

The result is therefore an application of Theorem 3.1(ii). \square

Remark 3.4 Corollary 3.3 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 3.3, we recover Proposition 1.3(ii).

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

Corollary 3.5 Suppose that \mathcal{X} is uniformly convex and uniformly smooth, let J be its duality mapping, set $q = \|\cdot\|^2/2$, and let $\varphi \in \Gamma_0(\mathcal{X})$. Then $q^* = \|\cdot\|_*^2/2$ and the following hold for every $x \in \mathcal{X}$.

- (i) $q(x) = (\varphi \square q)(x) + (\varphi^* \diamond q^*)(Jx)$.

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

Proof. This is an application of Theorem 3.1 with $f = q$. Indeed, $\text{dom } f = \mathcal{X}$, $\text{dom } f^* = \mathcal{X}^*$, and $\nabla f = J$. \square

In particular, if \mathcal{X} is a Hilbert space in Corollary 3.5, it follows from Remark 2.8(ii) and Remark 2.11 that we recover Moreau's decomposition principle (Proposition 1.3) and a fortiori Propositions 1.1 and 1.2. Another noteworthy instance of Corollary 3.5 is when $\varphi = \iota_K$, where K is a nonempty closed convex cone in \mathcal{X} . In this case, $\varphi^* = \iota_{K^\ominus}$, $\text{aprox}_\varphi^q = P_K$, and we derive from (1.4) and (1.5) that $\text{bprox}_\varphi^q = \Pi_K$. Hence, Corollary 3.5(ii)&(iii) yields Proposition 1.4.

Remark 3.6 Consider the setting of Theorem 3.1 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 2.5(iii), and (2.6) that we can rewrite Theorem 3.1(ii) as

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

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

Acknowledgement. The work of the first author was supported by the Agence Nationale de la Recherche under grant ANR-08-BLAN-0294-02. 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.), pp. 15–50. Marcel Dekker, New York, 1996.
- [2] Ya. I. Alber, Decomposition theorems in Banach spaces, in: *Operator Theory and Its Applications*, pp. 77–93. Fields Institute Communications, vol. 25, AMS, Providence, RI, 2000.
- [3] Ya. I. Alber, James orthogonality and orthogonal decompositions of Banach spaces, *Journal of Mathematical Analysis and Applications*, vol. 312, pp. 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*, pp. 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, *Communications in Contemporary Mathematics*, vol. 3, pp. 615–647, 2001.

- [6] H. H. Bauschke, J. M. Borwein, and P. L. Combettes, Bregman monotone optimization algorithms, *SIAM Journal on Control and Optimization*, vol. 42, pp. 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, *Journal of Convex Analysis*, vol. 8, pp. 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 and Control Letters*, vol. 55, pp. 45–51, 2006.
- [10] Y. Censor and S. A. Zenios, Proximal minimization algorithm with D-functions, *Journal of Optimization Theory and Applications*, vol. 73, pp. 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*, pp. 351–418. Ossolineum, Wrocław, 1979.
- [13] P. L. Combettes, Đinh Dũng, and B. C. Vũ, Dualization of signal recovery problems, *Set-Valued and Variational Analysis*, vol. 18, pp. 373–404, 2010.
- [14] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM Journal on Optimization*, vol. 18, pp. 1351–1376, 2007.
- [15] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Modeling and Simulation*, vol. 4, pp. 1168–1200, 2005.
- [16] S.-P. Han and O. L. Mangasarian, Conjugate cone characterization of positive definite and semidefinite matrices, *Linear Algebra and Its Applications*, vol. 56, pp. 89–103, 1984.
- [17] J.-B. Hiriart-Urruty and Ph. Plazanet, Moreau’s decomposition theorem revisited, *Annales de l’Institut Henri Poincaré – Analyse Non Linéaire*, vol. 6, pp. 325–338, 1989.
- [18] J.-B. Hiriart-Urruty and A. Seeger, A variational approach to copositive matrices, *SIAM Review*, vol. 52, pp. 593–629, 2010.
- [19] Y. H. Hu and W. Song, Weak sharp solutions for variational inequalities in Banach spaces, *Journal of Mathematical Analysis and Applications*, vol. 374, pp. 118–132, 2011.
- [20] C. Lescarret, Applications “prox” dans un espace de Banach, *Comptes Rendus de l’Académie des Sciences de Paris Série A*, vol. 265, pp. 676–678, 1967.
- [21] Y. Lucet, What shape is your conjugate? A survey of computational convex analysis and its applications, *SIAM Review*, vol. 52, pp. 505–542, 2010.
- [22] J.-J. Moreau, Décomposition orthogonale d’un espace hilbertien selon deux cônes mutuellement polaires, *Comptes Rendus de l’Académie des Sciences de Paris Série A*, vol. 255, pp. 238–240, 1962.
- [23] J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *Comptes Rendus de l’Académie des Sciences de Paris Série A*, vol. 255, pp. 2897–2899, 1962.
- [24] J.-J. Moreau, Sur la fonction polaire d’une fonction semi-continue supérieurement, *Comptes Rendus de l’Académie des Sciences de Paris Série A*, vol. 258, pp. 1128–1130, 1964.

- [25] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.
- [26] R. T. Rockafellar, Level sets and continuity of conjugate convex functions, *Transactions of the American Mathematical Society*, vol. 123, pp. 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: *Non-smooth Mechanics and Analysis*, pp. 3–12. Springer-Verlag, 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, *Journal of Inverse and Ill-Posed Problems*, vol. 16, pp. 479–506, 2008.
- [30] W. Song and Z. Cao, The generalized decomposition theorem in Banach spaces and its applications, *Journal of Approximation Theory*, vol. 129, pp. 167–181, 2004.
- [31] M. Teboulle, Entropic proximal mappings with applications to nonlinear programming, *Mathematics of Operations Research*, vol. 17, pp. 670–690, 1992.
- [32] D. Wexler, Prox-mappings associated with a pair of Legendre conjugate functions, *Revue Française d’Automatique, Informatique et Recherche Opérationnelle*, vol. 7, pp. 39–65, 1973.
- [33] C. Zălinescu, *Convex Analysis in General Vector Spaces*. World Scientific, River Edge, NJ, 2002.