J Glob Optim (2012) 52:831–842
DOI 10.1007/s10898-011-9714-1

Generalized projections onto convex sets

O. P. Ferreira · S. Z. Németh

Received: 13 September 2010 / Accepted: 24 March 2011 / Published online: 8 April 2011

Abstract This paper introduces the notion of projection onto a closed convex set associated with a convex function. Several properties of the usual projection are extended to this setting. In particular, a generalization of Moreau's decomposition theorem about projecting onto closed convex cones is given. Several examples of distances and the corresponding generalized projections associated to particular convex functions are presented.

Keywords Projection · Convex set · Convex cone · Convex function

1 Introduction

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set and $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then, the classical constrained convex optimization problem is

$$\min\{f(x) : x \in C\}.$$

In this paper, we deal with the properties of the solution set for particular cases of this problem. Let us introduce the particular problem which we are interested in. The projection onto a convex set is obtained by minimizing the distance function on the considered convex set. Specifically, given a distance function d and a nonempty closed convex set C in the Euclidean space $\mathbb{R}^n$, the projection of the point $x \in \mathbb{R}^n$ onto C with respect to d, is the set

This work was done while the first author was visiting the second author at the School of Mathematics in The University of Birmingham, during the period of January 2010 to July 2010.

O. P. Ferreira (✉)
IME/UFG, Campus II - Caixa Postal 131, Goiânia, GO 74001-970, Brazil
e-mail: orizon@mat.ufg.br

S. Z. Németh
School of Mathematics, The University of Birmingham, The Watson Building, Edgbaston, Birmingham B15 2TT, United Kingdom
e-mail: nemeths@for.mat.bham.ac.uk

$$P_C(x) := \{p \in C : d(x, p) \le d(x, y), \quad \forall\, y \in C\}; \tag{1}$$

that is, the minimizer set of the function $C \ni y \mapsto d(x, y)$. In this sense, for each fixed distance function we have an associated projection set. This view point allows the study of the notion of projection in several different contexts. Works dealing with projections associated to different distances include Carrizosa and Plastria [6], Censor and Elfving [7], Censor et al. [9], Mangasarian [19,20], Dax [11,12], Plastria and Carrizosa [27] and Scolnik et al. [30].

The interest in the subject of projection arises from several situations, having a wide range of applications in pure and applied mathematics such as Functional Analysis (see e.g. [33]), Convex Analysis (see e.g. [15]), Optimization (see e.g. [2,8,9,26,28,30,32]), Numerical Linear Algebra (see e.g. [31]), Statistics (see e.g. [4,13]), Computer Graphics (see e.g. [14]) and Ordered Vector Spaces (see e.g. [16,17,22,24,25]).

Several distance functions defined in Euclidean spaces are convex functions. For instance, the distances associated with the Euclidean norm $\|\cdot\|$ is defined by

$$d(x, y) = \|x - y\|, \quad x, y \in \mathbb{R}^n,$$

and therefore, the convexity of $d(x, \cdot) : \mathbb{R}^n \to \mathbb{R}$, the distance function from the point x, is a consequence of the positive homogeneity and triangle inequality property of the norm. Since the projection of the point $x \in \mathbb{R}^n$ onto a closed convex set C is the minimizer set of the distance function from the point x on the convex set C, see (1), it is natural to extend the concept of projection as the minimizer set of positive convex functions. Therefore, we study the minimizer set of the following problem:

$$\min\{\varphi(x - y) : y \in C\},$$

where $x \in \mathbb{R}^n$, $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a convex function and C is a nonempty closed convex set. This extension facilitates deriving several important properties of the projection associated to different distances in a unified manner.

Using this general approach, we will study properties of the projection onto a convex set associated with a convex function. We present several characterizations of the projection onto a convex set associated with a convex function, which extends the usual characterizations of Zarantonello for projections [33]. Although most of the results hold in general Hilbert spaces too, for simplicity of the ideas we will present them in Euclidean spaces only. In particular, we will extend the Moreau's theorem for projections onto convex cones (see [21]) to this more general setting.

It is worth to remark that in applications, several other distance-like notions and projections were considered (see e.g. proximity mappings, Bregman distance [5], Kullback–Leibler divergence, Csiszars f-divergence [10] , etc.) Most of these are also generalizations of the Euclidean distance. Our approach is different since it focuses on the generalization of the Euclidean norm (hence it is translation invariant in contrast to some of the above mentioned extensions). Of course, various questions occurs with respect to the behavior of these projections which constitute topics for the geometry of normed spaces, abstract best approximation theory (in the case of norms the history goes back to the nineteenth century), etc. Our view lies in the fact that we are searching for extensions of geometrical properties of Euclidean projections.

The structure of this paper is as follows. In Sect. 2 we define the distance and the projection with respect to a convex function. We also present several examples of convex functions, which generate distances and projections, along with establishment several properties of the distance and the projection onto a convex set and, in particular, onto a convex cone and onto a hyperplane. We conclude this paper by making some final remarks in Sect. 3.

2 Generalized projection

In this section we define the distance and the projection with respect to a convex function. Several examples of convex functions which generate distances and projections will be presented. Besides, we will establish several properties of the projection onto a convex set and, in particular, onto a convex cone and onto a hyperplane.

Throughout the paper we suppose that $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a convex function satisfying the following two conditions:

H1. $\varphi(0) = 0$;
H2. $\varphi(x) \geq 0, \quad \text{for all } x \in \mathbb{R}^n$.

We also consider the following two conditions on the function φ, which *will be considered to hold only when explicitly stated*:

H3. $\varphi(x) = \varphi(-x), \quad \text{for all } x \in \mathbb{R}^n$;
H4. $\varphi(\lambda x) = \lambda\varphi(x), \quad \text{for all } \lambda \geq 0$.

Remark 1 Rockafellar [29] called a convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfying conditions **H1**, **H2** and **H4** a gauge function, and these functions are characterized as

$$\varphi(x) = \inf\{t \geq 0 : x \in tB_\varphi\}, \quad B_\varphi = \{x \in \mathbb{R}^n : \varphi(x) \leq 1\}.$$

Note that condition **H3** together with **H4** is equivalent to $\varphi(\lambda x) = |\lambda|\varphi(x)$, for all $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. All convex functions φ satisfying **H4** are subadditive; that is, $\varphi(x+y) \leq \varphi(x)+\varphi(y)$, for all $x, y \in \mathbb{R}^n$ (see, Theorem 4.7 page 30 of [29]). Thus, if φ is positive everywere except at the origin and satisfies **H1**, **H2**, **H3** and **H4**, then φ is a norm (see, page 131 of [29]).

Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. The *distance function* $d^\varphi(\cdot, C) : \mathbb{R}^n \to \mathbb{R}$ *to* C *with respect to* φ is defined by

$$d^\varphi(x, C) := \inf\{\varphi(x - y) : y \in C\}, \tag{2}$$

and the *projection mapping* $P_C^\varphi(\cdot) : \mathbb{R}^n \multimap C$ *with respect to* φ *onto the set* C is defined by

$$P_C^\varphi(x) := \{p \in C : \varphi(x - p) \leq \varphi(x - y), \quad \forall\, y \in C\}. \tag{3}$$

Conditions **H1** and **H2** imply that φ has 0 as a minimizer. As a consequence, $z \in P_C^\varphi(z)$ for all $z \in C$. Using the previous two equalities, we conclude that

$$d^\varphi(x, C) = \varphi(x - p), \quad \forall\, p \in P_C^\varphi(x). \tag{4}$$

If the function $C \ni y \mapsto \varphi(x - y)$ has exactly one minimizer for each $x \in \mathbb{R}^n$ then the projection mapping P_C^φ is single valued. In this case, the last equality becomes

$$d^\varphi(x, C) = \varphi\left(x - P_C^\varphi(x)\right).$$

Remark 2 The projection mappings associated with the convex functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfying conditions **H1**, **H2**, **H2** and **H4** are in general multivalued. In general, the projection mapping associated with the 1-norm and the $+\infty$-norm onto a convex cone is not single valued. However, the projection mapping associated with the 1-norm onto the cone $\mathbb{R}^n_+$ is single valued, whereas the projection mapping associated with $+\infty$-norm onto the cone $\mathbb{R}^n_+$ is not.

Remark 3 A convex set C is strictly convex if its boundary bdC does not contain any line segment. Formally this means that for each $x, y \in bdC$ with $x \neq y$ there is no $0 < \lambda < 1$ such that $\lambda x + (1 - \lambda)y \in bdC$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *strongly quasiconvex* if for each $0 < \lambda < 1$ and each $x, y \in \mathbb{R}^n$ with $x \neq y$ we have

$$f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\}.$$

All strictly convex functions are strongly quasiconvex, but there exist strongly quasiconvex functions, which are not strictly convex (for example the Euclidean norm). It follows from the definition that the function φ is strongly quasiconvex if and only if all of its nonempty sublevel sets $\{x \in \mathbb{R}^n : \varphi(x) \leq L\}$ for $L \in \mathbb{R}$, are strictly convex (remember that φ is continuous). Moreover, for each $x \in \mathbb{R}^n$ the function $C \ni y \mapsto \varphi(x - y)$ has only one minimizer, see Theorem 3.5.9 of [3]. Consequently, the projection mapping associated with a strongly quasiconvex function is single valued.

Remark 4 Different functions can generate the same projection mapping. Indeed, taking $s > 1$ the function $\psi = \varphi^s$ is convex, satisfies **H1**, **H2** and **H3**. Moreover, we have

$$d^{\psi}(\cdot, C) = (d^{\varphi}(\cdot, C))^s, \quad P^{\psi}_C = P^{\varphi}_C.$$

As a consequence, the new optimization problems defining the distance functions, may become differentiable and therefore easier to handle. For instance, if a norm comes from a scalar product, then its square is differentiable everywhere and the projection with respect to the square of the norm is equal to the projection with respect to the norm.

2.1 Properties of the generalized projection

In this section we present some basic properties of the projection mapping with respect to a convex function. We begin by giving a characterization of the projection onto C with respect to φ.

Proposition 1 *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set, $x \in \mathbb{R}^n$ and $p \in C$. Then, the following statements are equivalent:*

(i) $p \in P^{\varphi}_C(x)$;
(ii) *there exists $u \in \partial\varphi(x - p)$ such that $\langle u, q - p\rangle \leq 0$, for all $q \in C$.*

Proof It is easy to see, that for each $x \in \mathbb{R}^n$ the function $\mathbb{R}^n \ni y \mapsto \varphi_x(y) = \varphi(x - y)$ is convex and $\partial\varphi_x(y) = -\partial\varphi(x - y)$. Hence, applying the optimality condition for the optimization problem in (2) (see Theorem 3.4.3 of [3]) and taking in account the definition of the projection in (3) the equivalence of items (i) and (ii) follows. □

Corollary 2 *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set, $x \in \mathbb{R}^n$ and $p \in C$. If φ is differentiable at $x - p$, then the following statements are equivalent:*

(i) $p \in P_C^{\varphi}(x)$;
(ii) $\langle \nabla\varphi(x-p), q-p\rangle \le 0$, *for all* $q \in C$.

Moreover, if the function φ *is differentiable at* $x-y$ *for every* $y \in C$ *then:*

(iii) $\langle \nabla\varphi(x-q), q-p\rangle \le 0$, *for all* $q \in C$,

is equivalent to items (i) and (ii).

Proof The equivalence of items (i) and (ii) follows directly from Proposition 1.

Since the function $\mathbb{R}^n \ni y \mapsto \varphi(x-y)$ is convex and we are assuming differentiability at $x-y$ for every $y \in C$, we conclude that the gradient is monotonous and continuous at $x-y$ for every $y \in C$ (see Theorem 25.5 on page 246 of [29]). Using the monotonicity and continuity of the gradient, the equivalence of items (ii) and (iii) follows from Lemma 1.5 of [18]. □

By using Proposition 1, in the next proposition, we present a property of the generalized projection mapping associated with generalized distances, extending Theorem 1.1 of [33]. This theorem is an important tool in spectral theory of metric projections (see [33]).

Proposition 3 *Let* $C \subset \mathbb{R}^n$ *be a nonempty closed convex set and* $P : \mathbb{R}^n \multimap C$ *be a surjective mapping. Then, the following statements are equivalent:*

(i) $P(x) \subset P_C^{\varphi}(x)$, *for all* $x \in \mathbb{R}^n$;
(ii) *for all* $x \in \mathbb{R}^n$ *and all* $p \in P(x)$, *there exists* $u \in \partial\varphi(x-p)$ *such that* $\langle u, q-p\rangle \le 0$, *for all* $q \in P(y)$ *and all* $y \in \mathbb{R}^n$.

Proof Let $x \in \mathbb{R}^n$ and $p \in P(x)$. Since $P(x) \subset P_C^{\varphi}(x)$ and $P(y) \subset C$ for all $y \in \mathbb{R}^n$, by using Proposition 1, we conclude that item (i) implies item (ii).

In order to prove that item (ii) implies item (i), take $x \in \mathbb{R}^n$ and $p \in P(x)$ arbitrarily. Since the mapping P is surjective, it follows that for each $q \in C$, there exists $y \in \mathbb{R}^n$, such that $q \in P(y)$. Thus, using item (ii), we conclude that there exists $u \in \partial\varphi(x-p)$ such that

$$\langle u, q-p\rangle \le 0, \quad \forall q \in C,$$

and using Proposition 1 we conclude that $p \in P_C^{\varphi}(x)$. Hence $P(x) \subset P_C^{\varphi}(x)$, for all $x \in \mathbb{R}^n$. Therefore, the statements are equivalent. □

Now, using Corollary 2 we present several equivalences related with the generalized projection mapping associated with generalized distances, extending Theorem 1.1, Lemma 1.2 and Theorem 1.2 of [33]. All these results are important tools in spectral theory of metric projections (see [33]). Therefore these results would be a good starting point for developing a spectral theory for more general projection mappings.

Proposition 4 *Let* $C \subset \mathbb{R}^n$ *be a nonempty closed convex set and* $P : \mathbb{R}^n \multimap C$ *a surjective mapping. If* φ *is differentiable in* $\mathbb{R}^n\backslash\{0\}$, *then the following two statements are equivalent.*

(i) $P(x) \subset P_C^{\varphi}(x)$, *for all* $x \in \mathbb{R}^n$.
(ii) $\langle \nabla\varphi(x-p), q-p\rangle \le 0$, *for all* $x, y \in \mathbb{R}^n$, *all* $p \in P(x)$ *with* $p \ne x$, *and all* $q \in P(y)$.

If $z \in P(z)$ *for all* $z \in C$ *and* φ *is differentiable in* $\mathbb{R}^n$, *then items (i) and (ii) are equivalent to the statements*

(iii) $\langle \nabla\varphi(x-p) - \nabla\varphi(y-q), p-q\rangle \geq 0$, *for all* $x, y \in \mathbb{R}^n$, *all* $p \in P(x)$, *and all* $q \in P(y)$.

(iv) $\nabla\varphi(P^{-1} - I)$ *is monotone, where the multivalued mapping* $\nabla\varphi(P^{-1} - I) : C \multimap \mathbb{R}^n$ *is defined by*

$$\nabla\varphi(P^{-1} - I)(p) := \{\nabla\varphi(x-p) : x \in \mathbb{R}^n, \quad p \in P(x)\}. \tag{5}$$

Furthermore, if $z \in P(z)$ *for all* $z \in C$ *and* φ *is differentiable in* $\mathbb{R}^n$, *then items (i), (ii), (iii) and (iv) are equivalent to*

(v) $\langle \nabla\varphi(x-q), q-p\rangle \leq 0$, *for all* $x, y \in \mathbb{R}^n$, *all* $p \in P(x)$, *and all* $q \in P(y)$.

Proof The equivalence of items (i) and (ii) follows from Proposition 3.
In order to prove that item (i) implies item (iii), take $x, y \in \mathbb{R}^n$, $p \in P(x)$ and $q \in P(y)$. Since $P(z) \subset P_C^{\varphi}(z) \subset C$, for all $z \in \mathbb{R}^n$, Corollary 2 gives

$$\langle \nabla\varphi(x-p), q-p\rangle \leq 0, \quad \langle \nabla\varphi(y-q), p-q\rangle \leq 0.$$

Using the last two inequalities, some simples algebraic manipulations yields

$$\langle \nabla\varphi(x-p) - \nabla\varphi(y-q), p-q\rangle \geq 0.$$

As the last inequality holds for all $x, y \in \mathbb{R}^n$, $p \in P(x)$ and $q \in P(y)$ the implication is proved. In order to that item (iii) implies item (i), take $x \in \mathbb{R}^n$ and $p \in P(x)$. Since we assume $q \in P(q)$ for all $q \in C$, taking $y = q$ in item (iii), we conclude that

$$\langle \nabla\varphi(x-p) - \nabla\varphi(0), p-q\rangle \geq 0, \quad \forall q \in C.$$

Since 0 is a minimizer of φ we have $\nabla\varphi(0) = 0$. Thus the above inequality implies

$$\langle \nabla\varphi(x-p), q-p\rangle \leq 0, \quad \forall q \in C.$$

Thus, using Corollary 2 we have $p \in P_C^{\varphi}(x)$ and the implication follows, in effect, when $z \in P(z)$ for all $z \in C$ items (iii) and (i) are equivalent.

Next we prove the equivalence of items (iii) and (iv). Preliminarily, note that the definition of the multivalued mapping $\nabla\varphi(P^{-1} - I)$ in (5) implies

$$\begin{aligned}&\nabla\varphi(P^{-1} - I)(p) - \nabla\varphi(P^{-1} - I)(q)\\ &= \{\nabla\varphi(x-p) - \nabla\varphi(y-q) : x, y \in \mathbb{R}^n, p \in P(x), q \in P(y)\}.\end{aligned}$$

Therefore, the equivalence is an immediate consequence of the definition of monotonicity and the above equality. Consequently, items (i), (ii), (iii) and (iv) are equivalent.

In order to conclude the proof, it suffices to show that item (i) and item **v** are equivalent. Take $x, y \in \mathbb{R}^n$, $p \in P(x)$ and $q \in P(y)$. Since $P(x) \subset P_C^{\varphi}(x)$ and $P(y) \subset C$ we have $p \in P_C^{\varphi}(x)$ and $q \in C$. Using that item (i) in Corollary 2 implies item (iii) of it, we conclude that

$$\langle \nabla\varphi(x-q), p-q\rangle \leq 0.$$

Since the last inequality holds for all $x \in \mathbb{R}^n$, $p \in P(x)$ and $q \in P(y)$, we obtain that item (i) implies item **v**. Conversely, take $x \in \mathbb{R}^n$, $p \in P(x)$. Since the mapping P is surjective, for each $q \in C$ there exists $y \in \mathbb{R}^n$ such that $q \in P(y)$. Hence, from item **v** it follows that

$$\langle \nabla\varphi(x-q), q-p\rangle \leq 0, \quad \forall q \in C.$$

Using that item (iii) in Corollary 2 implies item (i) of it, we obtain $p \in P_C^{\varphi}(x)$. So, $P(x) \subset P_C^{\varphi}(x)$, for all $x \in \mathbb{R}^n$, which prove our last implication. □

Remark 5 If P is single valued and surjective then the condition $z \in P(z)$ for all $z \in C$, in the above proposition, becomes $P^2 = P$.

2.2 Properties of the generalized projection onto a convex cone

In this section we present properties of the projection mapping onto a convex cone with respect to a convex function, in particular, we present a generalization of Moreau's theorem [21] for this new context.

A closed set $K \subset \mathbb{R}^n$ is called a *closed convex cone* if $\lambda x \in K$ and $x + y \in K$ for all $x, y \in K$ and $\lambda > 0$. Let $K \subset \mathbb{R}^n$ be a closed convex cone. The *polar cone* of K is the set

$$K^{\circ} := \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 0, \quad \forall\, y \in K\}.$$

The next result is a generalization of Moreau's theorem, see [21] (see also Theorem 3.2.5 of [15]).

Theorem 5 *Let $K \subset \mathbb{R}^n$ be a closed convex cone, $x \in \mathbb{R}^n$ and $p \in K$. Then, $p \in P_K^{\varphi}(x)$ if and only if there exists $u \in \partial\varphi(x - p)$ such that*

$$u \in K^{\circ}, \quad \langle u, p \rangle = 0.$$

Proof Let $x \in \mathbb{R}^n$ and $p \in P_K^{\varphi}(x)$. From Proposition 1 it follows that there exists $u \in \partial\varphi(x - p)$ such that

$$\langle u, q - p \rangle \leq 0, \quad \forall\, q \in K. \tag{6}$$

Since K is a cone, in the last inequality we can replace q with λq, where $\lambda > 0$ to obtain $\langle u, \lambda q - p \rangle \leq 0$, for all $q \in K$. Since $\lambda > 0$, the latter inequality is equivalent to $\langle u, q - p/\lambda \rangle \leq 0$, for all $q \in K$. Thus, letting $\lambda \to +\infty$ we obtain

$$\langle u, q \rangle \leq 0, \quad \forall\, q \in K,$$

which implies that $u \in K^{\circ}$. For proving $\langle u, p \rangle = 0$, we use (6) with $q = 0$ and $q = 2p$ together with the assumption that K is a cone.

Conversely, let $x \in \mathbb{R}^n$, $p \in K$, and $u \in \partial\varphi(x - p)$ such that $u \in K^{\circ}$, $\langle u, p \rangle = 0$. Since $u \in K^{\circ}$, $\langle u, q \rangle \leq 0$, for all $q \in K$, which together with $\langle u, p \rangle = 0$ imply

$$\langle u, q - p \rangle \leq 0, \quad \forall\, q \in K.$$

As $p \in K$ and $u \in \partial\varphi(x - p)$, the latter inequality and Proposition 1 imply that $p \in P_K^{\varphi}(x)$ and therefore the proof of the theorem is concluded. □

In particular, if $\eta = \|\cdot\|$ is the Euclidean norm and $\varphi = \eta^2$, then we obtain the following well known characterization of the projection mapping (see for example Proposition 3.2.3 on page 120 of [15]). $p = P_K^{\eta}(x) = P_K^{\varphi}(x)$ if and only if

$$p \in K, \quad x - p \in K^{\circ}, \quad \langle x - p, p \rangle = 0.$$

This characterization is equivalent (see Proposition 3.2.3 on page 120 and Theorem 3.2.5 on page 121 of [15]) to Moreau's decomposition theorem (see [21]):

Theorem (Moreau). *Let $K \subset \mathbb{R}^n$ be a closed convex cone and $\eta = \|\cdot\|$ the Euclidian norm. For $x, p, q \in \mathbb{R}^n$, the following statement are equivalent:*

(i) $x = p + q$, $p \in K$, $q \in K^\circ$ *and* $\langle p, q\rangle = 0$;
(ii) $p = P^\eta_K(x)$ *and* $q = P^\eta_{K^\circ}(x)$.

The next lemma is an immediate consequence of **H3**, **H4** and the definition of the subdifferential.

Lemma 6 *The following statements are true:*

(i) *If* φ *satisfies* **H3**, *then* $\partial\varphi(x) = -\partial\varphi(-x)$, *for all* $x \in \mathbb{R}^n$;
(ii) *If* φ *satisfies* **H4**, *then* $\partial\varphi(x) = \partial\varphi(\lambda x)$, *for all* $x \in \mathbb{R}^n$ *and* $\lambda > 0$.

Proposition 7 *Let* $K \subset \mathbb{R}^n$ *be a closed convex cone. If* φ *satisfies* **H4**, *then there holds:*

$$P^\varphi_K(\lambda x) = \lambda P^\varphi_K(x), \quad \forall x \in \mathbb{R}^n, \quad \lambda > 0.$$

Proof First of all note that $\lambda P^\varphi_K(x) = \{\lambda p : p \in P^\varphi_K(x)\}$. Let $x \in \mathbb{R}^n$, $p \in P^\varphi_K(x)$ and $\lambda > 0$. From Theorem 5 there exists $u \in \partial\varphi(x - p)$ such that

$$u \in K^\circ, \quad \langle u, p\rangle = 0.$$

Since $u \in \partial\varphi(x-p)$, Lemma 6, item (ii) implies that $u \in \partial\varphi(\lambda x - \lambda p)$ and by using the above equality we trivially obtain that $\langle u, \lambda p\rangle = 0$. Thus, as $u \in K^\circ$ and $\lambda p \in K$, Theorem 5 implies that $\lambda p \in P^\varphi_K(\lambda x)$, which in turn implies that $\lambda P^\varphi_K(x) = \{\lambda p : p \in P^\varphi_K(x)\} \subset P^\varphi_K(\lambda x)$.

By putting in the obtained relation $1/\lambda$ in place of λ and λx in place of x we obtain the converse inclusion. □

Proposition 8 *Let* $K \subset \mathbb{R}^n$ *be a closed convex cone. If* φ *satisfies* **H3**, *then the following equality holds:*

$$P^\varphi_K(-x) = -P^\varphi_{-K}(x), \quad \forall x \in \mathbb{R}^n.$$

Proof Let $x \in \mathbb{R}^n$ and $p \in P^\varphi_K(-x)$. It follows from Theorem 5 that there exists $u \in \partial\varphi(-x - p)$ such that

$$u \in K^\circ, \quad \langle u, p\rangle = 0.$$

Since $u \in \partial\varphi(-x - p)$, Lemma 6, item (i) implies that $u \in -\partial\varphi(x + p)$. Hence, $-u \in \partial\varphi(x - (-p))$ and using the above equality we obtain trivialy that $\langle -u, -p\rangle = 0$. As $(-K)^\circ = -K^\circ$ and $u \in K^\circ$ we have $-u \in (-K)^\circ$. Since $-u \in \partial\varphi(x-(-p))$, $-u \in (-K)^\circ$ and $\langle -u, -p\rangle = 0$,, it follows from Theorem 5 that $-p \in P^\varphi_{-K}(x)$; that is, $p \in -P^\varphi_{-K}(x)$, which in turn implies that $P^\varphi_K(-x) \subset -P^\varphi_{-K}(x)$.

Whereby by symmetry $P^\varphi_{-K}(x) \subset -P^\varphi_K(-x)$ which is quite the converse of the previous inclusion, and the proof is complete. □

2.3 Properties of the generalized projection onto a hyperplane

In this section we present some properties of the projection mapping onto a hyperplane with respect to a convex function. In particular, following the idea of Mangasarian in [19], we give an explicit form for the projection mapping onto a hyperplane with respect to a convex function satisfying conditions **H1**, **H2** and **H4**. Note that, in this section, we are *not* supposing that φ satisfies condition **H3**.

It is an easy exercise to show that the next proposition is a special case of Proposition 1.

Proposition 9 *Let $L \subset \mathbb{R}^n$ be a hyperplane, $x \in \mathbb{R}^n$ and $p \in L$. Then, $p \in P_L^{\varphi}(x)$ if and only if there exists $u \in \partial\varphi(x-p)$ such that*

$$\langle u, q-p\rangle = 0, \quad \forall q \in L.$$

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a nonconstant convex function satisfying the conditions **H1**, **H2** and **H4**. The *dual function* $\varphi^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *associated with* φ is defined by

$$\varphi^*(y) := \sup\{\langle y, x\rangle : \varphi(x) = 1\}. \tag{7}$$

As an immediate consequence of the above definition we obtain Cauchy-Schwarz's inequality

$$\langle x, y\rangle \le \varphi^*(y)\varphi(x), \quad x, y \in \mathbb{R}^n, \quad \varphi(x) \ne 0, \tag{8}$$

and that φ^* is a convex function which also satisfies conditions **H1**, **H2** and **H4**.

Remark 6 For a convex function φ satisfying conditions **H1**, **H2** and **H4** Rockafellar [29] called the function φ^* defined in (7) the polar function of φ and φ a gauge function. Note that, if φ is positive except in the origin, then φ^* is always finite (see [29]). In this case, we may omit the condition $\varphi(x) \ne 0$ in (8), that is, the inequality (8) always holds.

Proposition 10 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a convex function satisfying conditions* **H1**, **H2** *and* **H4**. *Let $a \in \mathbb{R}$ and $v \in \mathbb{R}^n$ such that $v \ne 0$. Consider the hyperplane*

$$H = \{y \in \mathbb{R}^n : \langle v, y\rangle = a\}.$$

Let $P_H^{\varphi}(x)$ be the projection with respect to φ of the point $x \in \mathbb{R}^n$ onto the hyperplane H. Assume that $x \notin H$ and $\varphi(x-y) \ne 0$, for all $y \in H$. If $\varphi(v) \ne 0$, $\varphi(-v) \ne 0$, $\varphi^*(-v) < +\infty$ and $\varphi^*(v) < +\infty$ then $P(x) \subset P_H^{\varphi}(x)$, where

$$P(x) := \begin{cases} \left\{x - \frac{\langle v,x\rangle - a}{\varphi^*(v)}\, w : \varphi^*(v) = \langle v, w\rangle,\ \varphi(w) = 1\right\}, & \langle v, x\rangle > a; \\ \left\{x - \frac{\langle v,x\rangle - a}{\varphi^*(-v)}\, w : \varphi^*(-v) = \langle v, w\rangle,\ \varphi(-w) = 1\right\}, & \langle v, x\rangle < a. \end{cases} \tag{9}$$

and, as a consequence, the distance with respect to φ from the point $x \in \mathbb{R}^n$ to the hyperplane H is

$$d^{\varphi}(x, H) := \inf\{\varphi(x-y) : y \in H\} = \begin{cases} \frac{\langle v,x\rangle - a}{\varphi^*(v)}, & \langle v, x\rangle > a; \\ \frac{a - \langle v,x\rangle}{\varphi^*(-v)}, & \langle v, x\rangle < a. \end{cases} \tag{10}$$

If $\varphi(v) = 0$ and $\varphi(-v) = 0$ then $x - [\langle v, x\rangle - a/\|v\|^2]\, v \in P_H^{\varphi}(x)$ and $d^{\varphi}(x, H) = 0$.

Proof First assume that $\varphi(v) \ne 0$, $\varphi(-v) \ne 0$, $\varphi^*(-v) < +\infty$ and $\varphi^*(v) < +\infty$. In this case, $\varphi(v/\varphi(v)) = 1$, $\varphi(-v/\varphi(-v)) = 1$ and the definition of φ^* implies that

$$0 < \|v\|^2/\varphi(v) = \langle v, v/\varphi(v)\rangle \le \varphi^*(v), \quad 0 < \|v\|^2/\varphi(-v) = \langle -v, -v/\varphi(-v)\rangle \le \varphi^*(-v).$$

Hence, the set $P(x)$ defined in (9) is well defined. If $P(x) = \varnothing$, then the statement is trivial. Thus, assume $x \in \mathbb{R}^n$ so that $P(x) \ne \varnothing$ and $\langle v, x\rangle > a$. Take $w \in \mathbb{R}^n$ such that $\varphi^*(v) = \langle v, w\rangle$ and $\varphi(w) = 1$. Since $\varphi^*(v) = \langle v, w\rangle$, we obtain that

$$\left\langle v,\ x - \frac{\langle v, x\rangle - a}{\varphi^*(v)}\, w \right\rangle = \langle v,\ x\rangle - \frac{\langle v, x\rangle - a}{\varphi^*(v)}\, \langle v, w\rangle = a,$$

which implies that the set $P(x)$ is contained in H. Using assumption **H4**, we have

$$\varphi\left(x-\left(x-\frac{\langle v,x\rangle-a}{\varphi^*(v)}\,w\right)\right)=\frac{\langle v,x\rangle-a}{\varphi^*(v)}\,\varphi(w). \tag{11}$$

Since $\langle v, y\rangle = a$, for all $y \in H$ and taking into account that $\varphi(w) = 1$ and $\varphi(x-y) \neq 0$ for all $y \in H$, the last equality together with Eq. (8) gives

$$\varphi\left(x-\left(x-\frac{\langle v,x\rangle-a}{\varphi^*(v)}\,w\right)\right)=\frac{\langle v,x-y\rangle}{\varphi^*(v)}\leq\varphi(x-y),\quad\forall\, y\in H,$$

and, as $P(x) \subset H$, we conclude that $x - [(\langle v, x\rangle - a)/\varphi^*(v)]w$ is a minimizer of the function $\mathbb{R}^n \ni y \mapsto \varphi(x-y)$ on the hyperplane H. Therefore, $P(x) \subset P_H^{\varphi}(x)$. Since $x - [(\langle v, x\rangle - a)/\varphi^*(v)]w$ is a minimizer of the function $\mathbb{R}^n \ni y \mapsto \varphi(x-y)$ on the hyperplane H, using $\varphi(w) = 1$, we obtain from the equality in (11) and the definition of the distance that (10) holds. Hence, the first equalities in (9) and in (10) are proved. The second equality in (9) and in (10) may be proved by using similar arguments. This concludes the proof of the first part of our proposition.

Now, assume that $\varphi(v) = 0$ and $\varphi(-v) = 0$. Hence, if $\langle v, x\rangle > a$ then using **H4**, we have

$$\varphi\left(x-\left(x-\frac{\langle v,x\rangle-a}{\|v\|^2}\,v\right)\right)=\frac{\langle v,x\rangle-a}{\|v\|^2}\varphi(v)=0, \tag{12}$$

and if $\langle v, x\rangle < a$ then using **H4** again, we have

$$\varphi\left(x-\left(x-\frac{\langle v,x\rangle-a}{\|v\|^2}\,v\right)\right)=\frac{\langle v,x\rangle-a}{\|v\|^2}\varphi(-v)=0, \tag{13}$$

and the conditions **H1** and **H2** imply that $x - [(\langle v, x\rangle - a)/\|v\|^2]v$ is a minimizer of the function $\mathbb{R}^n \ni y \mapsto \varphi(x-y)$. As $x - [(\langle v, x\rangle - a)/\|v\|^2]v \in H$ we conclude that $x - [(\langle v, x\rangle - a)/\|v\|^2]v \in P_H^{\varphi}(x)$. Since $x - [(\langle v, x\rangle - a)/\|v\|^2]v$ is a minimizer of the function $H \ni y \mapsto \varphi(x-y)$, using (12), (13) and the definition of the distance we obtain $d^{\varphi}(x, H) = 0$, which concludes the proof. □

For the function φ satisfying condition **H3** we have $\varphi(v) = \varphi(-v)$ and $\varphi^*(v) = \varphi^*(-v)$. Therefore, the projection and distance associated with φ has a simpler formula. In this case, last proposition becomes:

Corollary 11 *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a convex function satisfying conditions* **H1***,* **H2***,* **H3** *and* **H4**. *Let $a \in \mathbb{R}$ and $v \in \mathbb{R}^n$ such that $v \neq 0$. Consider the hyperplane*

$$H = \{y \in \mathbb{R}^n : \langle v, y\rangle = a\}.$$

Let $P_H^{\varphi}(x)$ be the projection with respect to φ of the point $x \in \mathbb{R}^n$ onto the hyperplane H. Assume that $x \notin H$ and $\varphi(x-y) \neq 0$, for all $y \in H$. If $\varphi(v) \neq 0$ and $\varphi^*(v) < +\infty$ then $P(x) \subset P_H^{\varphi}(x)$, where

$$P(x) := \left\{x-\frac{\langle v,x\rangle-a}{\varphi^*(v)}\,w : \varphi^*(v)=\langle v,w\rangle,\ \varphi(w)=1\right\}, \tag{14}$$

and, as a consequence, the distance with respect to φ from the point $x \in \mathbb{R}^n$ to the hyperplane H is

$$d^{\varphi}(x,H) := \inf\{\varphi(x-y) : y\in H\} = \frac{|\langle v,x\rangle-a|}{\varphi^*(v)}. \tag{15}$$

If $\varphi(v) = 0$ then $x - [\langle v, x\rangle - a/\|v\|^2]\,v \in P_H^{\varphi}(x)$ and $d^{\varphi}(x, H) = 0$.

Remark 7 If φ is positive except at the origin and satisfies **H1**, **H2**, **H3** and **H4** (that is, if φ is a norm) then for all $v \in \mathbb{R}^n$ with $v \neq 0$ we have $0 < \varphi(v)$ and $0 < \varphi^*(v) < +\infty$. In this case, if $x \notin H$ in Proposition 10, then $\varphi(x - y) \neq 0$, for all $y \in H$. Finally, note that if $x \in H$, then $P(x) = \{x\} \subset P_H^{\varphi}(x)$. Hence, the conclusion of the Proposition 10 holds for $x \in H$ as well.

Using the above proposition we re-obtain, as a particular case, several examples of distances and associated projections already analyzed in the literature. For instance, the oblique projections used by Arioli et al. [1], Scolnik et al. [30], Censor and Elfving [7] and Censor et al. [9] and the projection mapping associated with the 1-norm, $+\infty$-norm and p-norm onto a hyperplane used by Mangasarian [19,20]. For applications using gauge distances; that is, distances associated with gauge functions (see Remark 6), as well as distances associated with different norms see [6,27].

3 Final remarks

In Sect. 2 we introduced the notion of projection with respect to a convex function. This notion becomes the usual projection when the considered convex function is the Euclidean norm. Since several important properties of the usual projection come from its nonexpansivity (see [33]), it would be interesting to characterize the class of convex functions for which the associated projection is non-expansive.

This paper deals only with Convex Optimization, and only touches slightly more general fields (generalized convexity see Remark 3 and max–min systems see [32]). We expect that the results in this paper constitute a first step towards a more general setting. However, in general, when solving variational inequalities and complementarity problems, iterative methods involving projections may be used. Some of those solutions are global solutions of the corresponding optimization problem, see [23,26] and the references therein. We foresee further progress along these line in the nearby future.

Acknowledgments The authors express their gratitude to the referees for the many helpful comments which led to a substantially improved version of the paper. The author O. P. Ferreira was supported in part by FUNAPE/UFG, CNPq Grants 201112/2009-4, 475647/2006-8 and PRONEX–Optimization(FAPERJ/CNPq). The author S. Z. Németh was supported in part by the Hungarian Research Grant OTKA 60480

References

1. Arioli, M., Duff, I., Noailles, J., Ruiz, D.: A block projection method for sparse matrices. SIAM J. Sci. Stat. Comput. **13**(1), 47–70 (1992)
2. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**(3), 367–426 (1996)
3. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming, Theory and Algorithms. Wiley, New York (1979)
4. Berk, R., Marcus, R.: Dual cones, dual norms, and simultaneous inference for partially ordered means. J. Am. Stat. Assoc. **91**(433), 318–328 (1996)
5. Butnariu, D., Kassay, G.: A proximal-projection method for finding zeros of set-valued operators. SIAM J. Control Optim. **47**(4), 2096–2136 (2008)
6. Carrizosa, E., Plastria, F.: Optimal expected-distance separating halfspace. Math. Oper. Res. **33**(3), 662–677 (2008)
7. Censor, Y., Elfving, T.: Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem. SIAM J. Matrix Anal. Appl. **24**(1), 40–58 (electronic) (2002)
8. Censor, Y., Elfving, T., Herman, G.T., Nikazad, T.: On diagonally relaxed orthogonal projection methods. SIAM J. Sci. Comput. **30**(1), 473–504 (2007/08)

9. Censor, Y., Gordon, D., Gordon, R.: Component averaging: an efficient iterative parallel algorithm for large and sparse unstructured problems. Parallel Comput. **27**(6), 777–808 (2001)
10. Csiszár, I.: Generalized projections for non-negative functions. Acta Math. Hung. **68**(1-2), 161–186 (1995)
11. Dax, A.: The distance between two convex sets. Linear Algebra Appl. **416**(1), 184–213 (2006)
12. Dax, A., Sreedharan, V.P.: Theorems of the alternative and duality. J. Optim. Theory Appl. **94**(3), 561–590 (1997)
13. Dykstra, R.L.: An algorithm for restricted least squares regression. J. Am. Stat. Assoc. **78**(384), 837–842 (1983)
14. Foley, J.D., Dam, A.van , Feiner, S.K., Hughes, J.F.: Computer Graphics: Principles and Practice. Addison-Wesley systems programming series, Reading (1990)
15. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms: Fundamentals. I, volume 305 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (1990)
16. Isac, G., Németh, A.B.: Monotonicity of metric projections onto positive cones of ordered Euclidean spaces. Arch. Math. Basel **46**(6), 568–576 (1986)
17. Isac, G., Németh, A.B.: Isotone projection cones in Euclidean spaces. Ann. Sci. Math. Québec **16**(1), 35–52 (1992)
18. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and their Applications, volume 88 of Pure and Applied Mathematics. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York (1980)
19. Mangasarian, O.L.: Arbitrary-norm separating plane. Oper. Res. Lett. **24**(1–2), 15–23 (1999)
20. Mangasarian, O.L.: Polyhedral boundary projection. SIAM J. Optim. **9**(4), 1128–1134 (electronic, Dedicated to John E. Dennis, Jr., on his 60th birthday) (1999)
21. Moreau, J.-J.: Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires. C R Acad. Sci. Paris **255**, 238–240 (1962)
22. Németh, A.B., Németh, S.Z.: How to project onto an isotone projection cone. Linear Algebra Appl. **433**(1), 41–51 (2010)
23. Németh, S.Z.: Iterative methods for nonlinear complementarity problems on isotone projection cones. J. Math. Anal. Appl. **350**(1), 340–347 (2009)
24. Németh, S.Z.: Characterization of latticial cones in Hilbert spaces by isotonicity and generalized infimum. Acta Math. Hung. **127**(4), 376–390 (2010)
25. Németh, S.Z.: Isotone retraction cones in Hilbert spaces. Nonlinear Anal. **73**(2), 495–499 (2010)
26. Pardalos, P.M., Rassias, T.M, Khan, A.A (eds.): Nonlinear Analysis and Variational Problems, volume 35 of Springer Optimization and its Applications. Springer, New York (2010) (In honor of George Isac)
27. Plastria, F., Carrizosa, E.: Gauge distances and median hyperplanes. J. Optim. Theory Appl. **110**(1), 173–182 (2001)
28. Rami, M.A., Helmke, U., Moore, J.B.: A finite steps algorithm for solving convex feasibility problems. J. Global Optim. **38**(1), 143–160 (2007)
29. Rockafellar, R.T.: Convex Analysis. Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ (1970)
30. Scolnik, H.D., Echebest, N., Guardarucci, M.T., Vacchino, M.C.: Incomplete oblique projections for solving large inconsistent linear systems. Math. Program **111**(1-2, Ser. B), 273–300 (2008)
31. Stewart, G.W.: On the perturbation of pseudo-inverses, projections and linear least squares problems. SIAM Rev. **19**(4), 634–662 (1977)
32. Tao, Y., Liu, G.-P., Chen, W.: Globally optimal solutions of max-min systems. J. Global Optim. **39**(3), 347–363 (2007)
33. Zarantonello, E.H.: Projections on convex sets in Hilbert space and spectral theory. I. Projections on convex sets. In: Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971), pp. 237–341. Academic Press, New York (1971)