Nonlinear Functional Analysis and Applications
Vol. 15, No. 2 (2010), pp. 307-319

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

SOME REMARKS ON PROXIMAL POINT ALGORITHM IN SCALAR AND VECTORIAL CASES

Marius Durea[1] and Radu Strugariu[2]

[1]Faculty of Mathematics, Al. I. Cuza University
Bd. Carol I, nr. 11, 700506 - Iaşi, Romania
e-mail: durea@uaic.ro

[2]Department of Mathematics, Gh. Asachi University
Bd. Carol I, nr. 11, 700506 - Iaşi, Romania
e-mail: rstrugariu@yahoo.com

Abstract. In this note we present a proximal point algorithm for vector optimization. The method we use adapts to the vectorial case some classical results in the scalar optimization by means of some scalarization techniques.

1. Introduction

The proximal point algorithm is known to be one of the most important theoretical schemes to find zeroes of maximal monotone operators and, in particular, to minimize convex lower-semicontinuous scalar functions. Its relevance is (mostly) a theoretical one and the success of its implementation depends on the effectiveness of the methods used to solve the involved subproblems. Starting with the pioneering paper of Rockafellar [12], which clearly fix some existing ideas in the previous literature and gives much more insights on the potential of the inexact version of the algorithm when applies to optimization problems, an important literature has grown on possible extensions and generalizations of this algorithm (see, for example the survey paper [8] and the references therein).

Recently, some attention was focused also on the case of vectorial functions: see [7], [1], [2]. In that general situation, the sense of minima is understood by means of a partial order relation induced on the output space by a closed

[0]Received March 31, 2009. Revised August 12, 2009.
[0]2000 Mathematics Subject Classification: 90C29, 65K05.
[0]Keywords: Vector optimization, proximal point algorithm, weak minima, strong minima.

convex pointed cone. The main aim of this paper is to continue the investigation on this topic and to present in a vectorial setting a natural proximal point algorithm derived from the classical one via some scalarization methods.

The outline of the paper is as follows. In the second section we briefly discuss the main setting and properties of the scalar inexact proximal point algorithm. Moreover we point out a result concerning the finite termination of this algorithm in infinite dimensional setting. A proximal-like algorithm in a nonconvex framework is given as well. The third section is devoted to the vectorial case. We introduce an inexact proximal point algorithm for vector optimization problems, we show that it has the same main features as the classical algorithm and we compares our results with other results in literature. In our theoretical tour we deal mainly with the weak minimum points, but we briefly consider as well the case of Pareto minimum points in the situation when the interior of the ordering cone is empty.

2. Scalar case

In this section we briefly remind some well-known facts about proximal point algorithm separately in a convex and a nonconvex framework. In both cases we propose as well some remarks on the use of this algorithm.

2.1. **Convex framework.**

Let X be a real Hilbert space and denote by $\langle \cdot, \cdot \rangle$ its scalar product. We shall tacitly identify the topological dual of X with the space X itself. The classical exact proximal point algorithm (EPPA, for short) for finding a minimum of a given convex lower-semicontinuous proper function $f : X \to \mathbb{R} \cup \{+\infty\}$ consists of generating a sequence $(x_n) \subset X$ by solving the convex program associated to the Moreau-Yosida regularization of the initial function:

$$x_{n+1} = \operatorname{argmin}_{x \in X} \left(f(x) + \frac{\lambda_n}{2} \|x - x_n\|^2 \right) \tag{2.1}$$

where (λ_n) is a bounded sequence of positive real numbers and x_0 is arbitrarily chosen. In fact, using the subdifferential calculus one gets that:

$$\lambda_n(x_n - x_{n+1}) \in \partial f(x_{n+1}),$$

where ∂ denotes the Fenchel subdifferential of a convex function. Then, it is shown in [11], [12] that in the case where f admits a minimum then (x_n) is bounded and weakly converges towards such a minimum. Subsequently, an important literature have been developed by generalizing this initial method of finding minima. In general, these generalizations concern the following aspects (see [8], [3] and the references therein):

- take x_{n+1} not as an exact solution of the problem

$$\min\left(f(x)+\frac{\lambda_n}{2}\|x_n-x\|^2\right) \tag{2.2}$$

  but as an approximate solution of it (inexact proximal point algorithm, IPPA, for short);
- consider another types of perturbations for f instead of the square of the norm (generalized proximal point algorithms);
- consider the case where f is not convex but a lower semicontinuous function (of course, in this case one should take $\in$ instead of $=$ in relation (2.1)) (nonconvex proximal point algorithms).

Let us remind the main features of the IPPA in the Rockafellar's sense ([12]). In this case, the iterate x_{n+1} is taken as an approximate solution of the regularization program in relation (2.2). Let (ε_n) be a sequence of positive real numbers s.t. $\sum_{n=0}^{\infty}\varepsilon_n<\infty$, consider $x_0\in X$ arbitrarily, and denote at every step by $\tilde{x}_{n+1}$ the unique solution of the problem (2.2), and take $x_{n+1}\in X$ s.t. $\|x_{n+1}-\tilde{x}_{n+1}\|<\varepsilon_n$. Then the sequence (x_n) is bounded if and only if f has a minimum point over X and in this case (x_n) weakly converges to a minimum point. Moreover,

$$\lim_{n\to\infty}\|x_{n+1}-x_n\|=\lim_{n\to\infty}\|\tilde{x}_{n+1}-x_n\|=0.$$

Remind that one says that $\overline{x}\in X$ is a sharp minimum of f if there exists $\mu>0$ s.t. for every $x\in X$

$$f(x)-f(\overline{x})\geq\mu\|x-\overline{x}\|.$$

In the case of EPPA (i.e. $x_{n+1}=\tilde{x}_{n+1}$) it is known that if X has finite dimension and if f admits a sharp minimum, then this algorithm terminates in a finite number of iterations (see [12, Proposition 8], [5, Theorem 6]). We observe here that this result can be extended to the infinite dimensional case and for IPPA in a sense we make precise below.

Proposition 2.1. *Suppose that f admits a sharp minimum point $\overline{x}$. Then the sequences of exact iterations $(\tilde{x}_n)_{n\geq 1}$ is stationary equal to $\overline{x}$ for n large enough.*

Proof. Since $\overline{x}$ is a sharp minimum point for f, taking into account [13, Theorem 3.10.1, Implication $(i)\Rightarrow(x)$] one has that

$$\mu U_X\subset\partial f(\overline{x}),$$

where U_X stands for the closed unit ball in X. Then one can show as in [5, Lemma 5] that for any $z \neq \overline{x}$ and $w \in \partial f(z)$, $\|w\| \geq \mu$. Indeed, one has that

$$\mu \|z - \overline{x}\|^{-1} (z - \overline{x}) \in \partial f(\overline{x})$$

whence

$$\left\langle w - \mu \|z - \overline{x}\|^{-1} (z - \overline{x}), z - \overline{x} \right\rangle \geq 0.$$

Therefore,

$$\mu \|z - \overline{x}\| \leq \langle w, z - \overline{x} \rangle \leq \|w\| \|z - \overline{x}\|,$$

and the claim is proved. Now, from the construction of the exact iterates, one has that

$$\lambda_n(x_n - \tilde{x}_{n+1}) \in \partial f(\tilde{x}_{n+1}).$$

The fact that $\lim_{n\to\infty} \|\tilde{x}_{n+1} - x_n\| = 0$, the boundedness of (λ_n) and the above claim assures that $\tilde{x}_{n+1} = \overline{x}$ for every n large enough. $\square$

2.2. **Nonconvex framework.**

Let us consider now the case of a nonconvex function. In this subsection, X is a reflexive Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ is a weakly lower semicontinuous function having bounded (whence weakly compact) level sets. In particular, this ensures the existence of a minimum point for f over X. A necessary and sufficient condition for boundedness of the level sets is the coercivity of f, i.e. $\lim_{\|x\|\to\infty} f(x) = +\infty$. In [3, Section 3] a generalized exact proximal point algorithm for f is studied. In this paper we consider, on similar lines, an IPPA under some minimal assumptions. First, consider a function $\psi : X \times X \to \mathbb{R}_+$ which is weakly continuous with respect to the first variable and weakly lower semicontinuous with respect to the second variable. Moreover, suppose that $\psi(x, x) = 0$ for every $x \in X$. For several methods to construct functions with these properties see, for example, [8]. Consider a bounded sequence (λ_n) of positive real numbers, a sequence (ε_n) of positive real numbers s.t. $\sum_{n=0}^{\infty} \varepsilon_n := C < \infty$ and an element $x_0 \in X$; denote:

$$A_n := \operatorname{argmin}_{x\in X} (f(x) + \lambda_n \psi(x_n, x)).$$

Note that in our assumptions $A_n \neq \emptyset$ because $f(\cdot) + \lambda_n \psi(x_n, \cdot)$ is weakly lower semicontinuous and the boundedness of the level sets of f and the positivity of ψ ensures the boundedness of the level sets of $f(\cdot) + \lambda_n \psi(x_n, \cdot)$. Take $x_{n+1} \in X$ s.t. $d(x_{n+1}, A_n) < \varepsilon_n$ (where, as usual, $d(x, A) := \inf\{\|x - a\| \mid a \in A\}$ denotes the distance between $x \in X$ and $A \subset X$). This means that there exists an element $\tilde{x}_{n+1} \in A_n$ s.t. $\|x_{n+1} - \tilde{x}_{n+1}\| < \varepsilon_n$. A convergence result for this algorithm is stated below.

Proposition 2.2. *Suppose, in addition, that f is a Lipschitz function (whence finite valued) and $\lambda_n \to 0$. Then the sequence (x_n) generated as above is bounded and every accumulation point of it is a global minimizer of f.*

Proof. Observe first that for every $n \geq 1$ one has:

$$f(\tilde{x}_{n+1}) + \lambda_n \psi(x_n, \tilde{x}_{n+1}) \leq f(x_n),$$

whence

$$f(x_{n+1}) - f(x_n) \leq -\lambda_n \psi(x_n, \tilde{x}_{n+1}) + L \left\| \tilde{x}_{n+1} - x_{n+1} \right\|,$$

where $L > 0$ is the Lipschitz constant of f. In particular,

$$f(x_{n+1}) - f(x_n) \leq L\varepsilon_n.$$

This inequality entails that

$$f(x_{n+1}) - f(x_0) \leq L \sum_{k=0}^{n} \varepsilon_k \leq LC.$$

Consequently $(x_n) \subset \{x \in X \mid f(x) \leq f(x_0) + LC\}$. Since f is supposed to have bounded level sets, one deduces that (x_n) is bounded as well.

Take now $\overline{x}$ as a weak accumulation point of (x_n). Note that such a point exists by the fact that the space is reflexive. Then there exists a strictly increasing sequence of natural numbers (n_k) such that x_{n_k} converges weakly to $\overline{x}$. The weak lower semicontinuity of f yields

$$f(\overline{x}) \leq \liminf_{k \to \infty} f(x_{n_k}).$$

As above, for every k and every $x \in X$,

$$f(\tilde{x}_{n_k}) + \lambda_{n_k - 1} \psi(x_{n_k - 1}, \tilde{x}_{n_k}) \leq f(x) + \lambda_{n_k - 1} \psi(x_{n_k - 1}, x),$$

whence

$$f(\tilde{x}_{n_k}) \leq f(x) - \lambda_{n_k - 1} \psi(x_{n_k - 1}, \tilde{x}_{n_k}) + \lambda_{n_k - 1} \psi(x_{n_k - 1}, x).$$

Since

$$f(x_{n_k}) - f(\tilde{x}_{n_k}) \leq L\varepsilon_{n_k - 1},$$

one deduces that

$$f(x_{n_k}) \leq L\varepsilon_{n_k - 1} + f(x) - \lambda_{n_k - 1} \psi(x_{n_k - 1}, \tilde{x}_{n_k}) + \lambda_{n_k - 1} \psi(x_{n_k - 1}, x).$$

We conclude that

$$f(x_{n_k}) \leq f(x) + L\varepsilon_{n_k - 1} + \lambda_{n_k - 1} \psi(x_{n_k - 1}, x).$$

We pass now to the lim inf when k goes to $+\infty$ and we take into account that ψ is weakly continuous in the first variable and the sequences (λ_n) and (ε_n) converge towards 0. Then we have:

$$f(\overline{x}) \leq f(x)$$

and since x is arbitrary in X, we get the conclusion. $\square$

3. Vectorial case

The goal of this section is to introduce and study, by use of a simple but effective scalarization technique, some proximal point schemes for minimizing vector-valued maps. Let X be a Hilbert space and Y be a Banach space. We consider a pointed closed convex proper cone $K \subset Y$ which introduces a partial order on Y by the equivalence $y_1 \leq_K y_2 \Leftrightarrow y_2 - y_1 \in K$.

We denote by $D(x, \varepsilon)$ the closed ball with center x and radius $\varepsilon > 0$. In order to cover the case of functions with real extended values we add to Y two abstract and distinct elements $-\infty_K, +\infty_K$ which do not belong to Y and we set $\overline{Y} := Y \cup \{-\infty_K, +\infty_K\}$. The ordering and algebraic rules we need for the new elements are the following ones: $-(+\infty_K) = -\infty_K$; $\forall y \in \overline{Y}, -\infty_K \leq_K y \leq_K +\infty_K, y + (+\infty_K) = (+\infty_K) + y = +\infty_K$; $\forall \lambda \in \mathbb{R}_+ := [0, \infty), \lambda \cdot (+\infty_K) = +\infty_K$.

We need now a topology on the new space $\overline{Y}$. We say that a set $V \subset \overline{Y}$ is a neighborhood of $+\infty_K$ if there exists $\theta > 0$ s.t. $(K \setminus D(0, \theta)) \cup \{+\infty_K\} \subset V$. Similarly, a set $V \subset \overline{Y}$ is a neighborhood of $-\infty_K$ if there exists $\theta > 0$ s.t. $(-K \setminus D(0, \theta)) \cup \{-\infty_K\} \subset V$. We denote by τ the norm topology on Y and we endow the space $\overline{Y}$ with the following topology:

$$\overline{\tau} = \tau \cup \{D \subset \overline{Y} \mid D \textit{ neighborhood for } -\infty_K \textit{ or } +\infty_K \textit{ and } D \setminus \{-\infty_k, +\infty_K\} \in \tau\}.$$

Every topological notion used on $\overline{Y}$ will be considered with respect to this topology. Let us consider a vectorial map f from X into $Y \cup \{+\infty_K\}$ which is proper, i.e. nonidentical equal $+\infty_K$. We set $\operatorname{dom} f := \{x \in X \mid f(x) \in Y\}$, which is nonempty because f is proper.

We address the following minimization problem:

$$\min_{x \in X} f(x). \tag{3.1}$$

The minimizers are considered with respect to the partial order relation introduced in Y by the cone K (see [10]). In this way, we say that $\overline{x} \in X$ is a Pareto minimizer of (3.1) if $f(\overline{x})$ is a minimal point for $f(X)$ with respect to $\leq_K$, i.e. $f(\overline{x}) \in Y$ and $(f(X) - f(\overline{x})) \cap -K = \{0\}$. We denote the topological interior of K by $\operatorname{int} K$ and in the case where $\operatorname{int} K \neq \emptyset$, we say that $\overline{x}$ is a weak minimizer of (3.1) if $f(\overline{x})$ is a weak minimal point for $f(X)$ with respect to $\leq_K$, i.e. $f(\overline{x}) \in Y$ and $(f(X) - f(\overline{x})) \cap -\operatorname{int} K = \emptyset$. It is a well-known fact that, in the vectorial setting, having $\operatorname{int} K \neq \emptyset$ is an important advantage, this condition being very important in a theoretical discussion. But this is quite restrictive because for many important particular Banach spaces, as

$\mathcal{L}^p$ $(1 < p < \infty)$ for example, the natural ordering cones have empty interiors. However, the case of cones with nonempty interior should be considered because it still corresponds to some infinite dimensional Banach spaces (as $C[0,1], \mathcal{L}^\infty$ for example) and to the finite dimensional spaces as well. For the results below we need to suppose that K has nonempty interior if it is not stated otherwise. At the end of our theoretical tour we shall briefly consider the case where $\operatorname{int} K = \emptyset$.

In order to adapt the proximal point algorithm we use a well-known scalarization functional (see [6, Section 2.3]). We present this functional in our specific case and for this let us fix $e \in \operatorname{int} K$. We work with the convention that $0 \cdot (+\infty) = +\infty$.

Theorem 3.1. *Define the functional* $\varphi_e : Y \cup \{+\infty_K\} \to \mathbb{R} \cup \{+\infty\}$ *as*

$$\varphi_e(y) = \begin{cases} \inf\{\lambda \in \mathbb{R} \mid y \in \lambda e - K\}, & \text{if } y \in Y \\ +\infty, & \text{if } y = +\infty_K. \end{cases}$$

This map has finite values on Y, *is sublinear on* $Y \cup \{+\infty_K\}$ *(hence convex), strictly monotone with respect to* $\operatorname{int} K$, *monotone with respect to* K, *Lipschitz on* Y *and for every* $\lambda \in \mathbb{R}$

$$\{y \in Y \mid \varphi_e(y) \le \lambda\} = \lambda e - K, \quad \{y \in Y \mid \varphi_e(y) < \lambda\} = \lambda e - \operatorname{int} K.$$

The next lemma links the minima of the composite function $\varphi_e \circ f$ with the minimizers of f.

Lemma 3.2. *If* $\overline{x} \in X$ *is a minimum of* $\varphi_e \circ f$ *then it is a weak minimizer of* f *as well.*

Proof. Since $\overline{x}$ is a minimum of $\varphi_e \circ f$ and f is nonidentical equal with $+\infty_K$ one deduces that $(\varphi_e \circ f)(\overline{x}) \in \mathbb{R}$ whence $f(\overline{x}) \in Y$. For every $x \in X$, we have

$$0 \le (\varphi_e \circ f)(x) - (\varphi_e \circ f)(\overline{x}),$$

i.e.

$$0 \le \varphi_e(f(x)) - \varphi_e(f(\overline{x})).$$

But, since φ_e is sublinear,

$$\varphi_e(f(x)) - \varphi_e(f(\overline{x})) \le \varphi_e(f(x) - f(\overline{x})),$$

whence

$$0 \le \varphi_e(f(x) - f(\overline{x}))$$

for every $x \in X$. The form of the level sets of φ_e from Theorem 3.1 ensures that

$$f(x) - f(\overline{x}) \notin -\operatorname{int} K$$

for every $x \in X$, whence the conclusion. □

First, we introduce a counterpart of the IPPA for the vectorial convex case. Consider the assumptions:

(A1): f is $K-$convex, i.e. for every $x_1, x_2 \in \operatorname{dom} f$, and every $\alpha \in (0,1)$

$$\alpha f(x_1) + (1-\alpha) f(x_2) - f(\alpha x_1 + (1-\alpha) x_2) \in K;$$

(A2): f has closed level sets with respect to the linear subspace generated by e, i.e. the set

$$\{x \in X \mid f(x) \leq_K \lambda e\}$$

is closed for every $\lambda \in \mathbb{R}$.

Consider now the proximal point algorithm for $\varphi_e \circ f : X \to \mathbb{R} \cup \{+\infty\}$. Take a bounded sequence of positive real numbers (λ_n), let (ε_n) be a sequence of positive real numbers s.t. $\sum_{n=0}^{\infty} \varepsilon_n < \infty$, consider $x_0 \in X$ arbitrarily, and at every step take $\tilde{x}_{n+1}$ as

$$\tilde{x}_{n+1} = \operatorname{argmin}_{x \in X} \left((\varphi_e \circ f)(x) + \frac{\lambda_n}{2} \|x_n - x\|^2 \right) \quad (3.2)$$

$$= \operatorname{argmin}_{x \in X} \left(\varphi_e (f(x) + \frac{\lambda_n}{2} \|x_n - x\|^2 e) \right). \quad (3.3)$$

Subsequently consider $x_{n+1} \in X$ s.t. $\|x_{n+1} - \tilde{x}_{n+1}\| < \varepsilon_n$. The following result shows that the properties of φ_e displayed in Theorem 3.1 are good enough to ensure that the main features of IPPA are preserved by the above algorithm.

Theorem 3.3. *Suppose that the assumptions* $(A1)$ *and* $(A2)$ *hold. Then the sequence generated by (3.2) is well-defined. Also, it is bounded if and only if the function* $\varphi_e \circ f$ *admits a minimum point. In this case,* (x_n) *is weakly convergent to a weak minimizer of* f.

Proof. We show that the scalar function $\varphi_e \circ f$ fulfills the conditions in the IPPA. First we show the convexity: take $x_1, x_2 \in X$ and $\alpha \in (0,1)$. If $f(x_1)$ or $f(x_2)$ equals $+\infty_K$, there is nothing to prove. Otherwise, by the definition of φ_e, for every $\varepsilon > 0$ one has

$$f(x_1) \in [\varphi_e(f(x_1)) + \varepsilon] e - K$$

and

$$f(x_2) \in [\varphi_e(f(x_2)) + \varepsilon] e - K,$$

whence

$$\alpha f(x_1) + (1-\alpha) f(x_2) \in [\alpha (\varphi_e \circ f)(x_1) + (1-\alpha)(\varphi_e \circ f)(x_2) + \varepsilon] e - K.$$

Therefore, from $(A1)$,

$$f(\alpha x_1 + (1-\alpha)x_2) \in \alpha f(x_1) + (1-\alpha) f(x_2) - K$$
$$\subset [\alpha(\varphi_e \circ f)(x_1) + (1-\alpha)(\varphi_e \circ f)(x_2) + \varepsilon]e - K.$$

This shows that

$$\varphi_e(f(\alpha x_1 + (1-\alpha)x_2)) \leq \alpha(\varphi_e \circ f)(x_1) + (1-\alpha)(\varphi_e \circ f)(x_2) + \varepsilon$$

and letting $\varepsilon \to 0$ we obtain the desired property. Moreover for any $\lambda \in \mathbb{R}$ one has

$$\{x \in X \mid (\varphi_e \circ f)(x) \leq \lambda\} = \{x \in X \mid f(x) \leq_K \lambda e\}$$

whence, by $(A2)$, the level sets of $\varphi_e \circ f$ are closed and then $\varphi_e \circ f$ is a convex lower semicontinuous proper function.

We can apply now the conclusions of the classical IPPA to conclude that the sequence (x_n) generated by (3.2) is bounded if and only if $\varphi_e \circ f$ admits a minimum point. Also, in such a case (x_n) is weakly convergent to one of the minima of $\varphi_e \circ f$, say $\overline{x}$. Lemma 3.2 shows that $\overline{x}$ is a weak minimizer of f, whence the conclusion. □

If we suppose, in addition, that f has bounded level sets with respect to the linear subspace generated by e, we still have that (x_n) generated by (3.2) is well-defined, bounded and weakly convergent to a weak minimizer of f. This is true because in such a case the level sets of $\varphi_e \circ f$ are bounded and closed, whence weakly compact, because X is in particular a reflexive Banach space. Since $\varphi_e \circ f$ is a proper convex lower semicontinuous function, it admits a minimum point and the conclusion follows. Also, the fact that the level sets with respect to the linear subspace generated by e are bounded is equivalent with the boundedness of all the level sets of f, as an easy calculus shows.

Let us compare the above result with some existing results in literature. In [6, Section 4.2] the authors propose a similar scalarization technique to construct a generalized proximal point algorithm for vector equilibrium problems. However, they work on finite dimensional spaces and with an exact version of the algorithm. In [1] the authors proposed an algorithm similar with our construction. The differences are that the algorithm in [1] is more flexible in a sense (it allows to choose x_{n+1} among the weak minimal points of $f(\cdot) + \frac{\lambda_n}{2} \|x_n - \cdot\|^2 e_n$ where $(e_n) \subset \operatorname{int} K$) and, in contrast, our algorithm, which is essentially a scalar one, works under some different assumptions and requires at every step the calculation of a solution of a scalar problem which is technically easier to be done than the calculation of a weak minimum point of a vectorial program. Let us denote by $\langle \cdot, \cdot \rangle$ the canonical duality pairing between Y and its dual Y^* and by K^* the dual polar cone of K. Then remark that for any $y \in Y$, $\{x \in X \mid f(x) \leq_K y\} = \bigcap_{y^* \in K^*} \{x \in X \mid (y^* \circ f)(x) \leq \langle y, y^* \rangle\}$, hence if f is supposed (as in [1]) to be positive lower semicontinuous, i.e. $y^* \circ f$ is lower semicontinuous for every $y^* \in K^*$, then f has closed level sets, whence $(A2)$ is fulfilled. Moreover, the scalarization technique we use here could successfully replace the scalarization functional presented in [7] in order to derive a steepest descent method for vector optimization problems. Finally, let us quote here as well the paper [2] where the authors consider some inexact vectorial algorithms in some very similar conditions, statements and proofs as in [1].

The notion of sharp minimum for vector-valued functions have been defined (see, e.g., [9]) by means of the oriented distance function. We remind that if $y \in Y$ and $M \subset Y$, the oriented distance function is defined as $\Delta(y, M) := d(y, M) - d(y, Y \setminus M)$. A point $\overline{x} \in X$ is termed sharp minimum for the vector-valued function $f : X \to Y \cup \{+\infty_K\}$ if $\overline{x} \in \operatorname{dom} f$ and there exists $\gamma > 0$ s.t. for every $x \in \operatorname{dom} f$, $\Delta(f(x) - f(\overline{x}), -K) \geq \gamma \|x - \overline{x}\|$. In [4] it is shown that $\overline{x}$ is a sharp minimum for f if and only if there exists $\mu > 0$ s.t. for every $x \in \operatorname{dom} f$

$$\varphi_e(f(x) - f(\overline{x})) \geq \mu \|x - \overline{x}\|. \tag{3.4}$$

From this characterization one can deduce that $\overline{x}$ is a sharp minimum for the vectorial function f if and only if it is a sharp minimum for the scalar function $\varphi_e(f(\cdot) - f(\overline{x}))$ (note that $\varphi_e(0) = 0$). Since, in general,

$$\varphi_e(f(x) - f(\overline{x})) \geq (\varphi_e \circ f)(x) - (\varphi_e \circ f)(\overline{x})$$

one observes that, unfortunately, the above mentioned notion of sharp minimum in the vectorial setting which do not ensures the uniqueness of the minimum, is not enough to ensure the finite termination of algorithm (3.2). For this reason we introduce here a stronger notion in order to fulfill the desired task. One says that $\overline{x}$ is a strong sharp minimum for f if is sharp minimum for the scalar function $\varphi_e \circ f$. In other words, one says that $\overline{x}$ is a strong sharp minimum for f if there exists $\mu > 0$ s.t. for every $x \in \operatorname{dom} f$

$$(\varphi_e \circ f)(x) - (\varphi_e \circ f)(\overline{x}) \geq \mu \|x - \overline{x}\|. \tag{3.5}$$

In particular, the relation (3.5) ensures the uniqueness of the strong sharp minimum. Observe that both sharp and strong sharp concepts cover the notion of sharp minimum in the scalar case, because if $X = \mathbb{R}$, $K = \mathbb{R}_+$ and $e = 1$, then $\varphi_e(y) = y$ for any $y \in \mathbb{R}$. Consequently, one has the following result.

Theorem 3.4. *Suppose that f admits a strong sharp minimum. Then, under assumptions $(A1)$ and $(A2)$, the recursion (3.2) reaches the strong sharp minimum after a finite number of iterates.*

Proof. As in the previous theorem, the function $\varphi_e \circ f$ is convex and since it admits a sharp minimum, one can apply the mentioned fact about finite termination of PPA in the scalar case. □

We end the convex case with a brief discussion on the practical possibility to apply the vectorial algorithm (3.2) introduced by means of the functional φ_e. In our opinion, this algorithm is a very natural one because, if compares with the classical scalar case, it supposes, in addition, to solve at every step a rather easy (in view of the properties of φ_e) scalar problem. For example, observe that if $Y = \mathbb{R}^q$ ($q \in \mathbb{N}, q \geq 1$), the natural choice for the cone K is the usual ordering convex cone $\mathbb{R}^q_+ := \{(x^1, x^2, ..., x^q) \in \mathbb{R}^q \mid x^i \geq 0, \forall i \in \overline{1,q}\}$ and in this case, for $e = (1, 1, ..., 1)$, one has:

$$\varphi_e(y) = \max(y^1, y^2, ..., y^q).$$

Then the vectorial algorithm has the same computational properties as the scalar algorithm. Moreover for $q = 1$ one gets exactly the scalar algorithm.

Now we point out how the same technique of scalarization can be used to adapt to the vectorial case the scalar algorithm in Subsection 2.2. Let X be a reflexive Banach space and $f : X \to Y$ be a vectorial Lipschitz function. To adapt the previous result concerning nonconvex case, we need another assumption, namely:

(A2'): f has bounded and weakly closed level sets with respect to the linear subspace generated by e (i.e. the set

$$\{x \in X \mid f(x) \leq_K \lambda e\}$$

is closed bounded and weakly closed for every $\lambda \in \mathbb{R}$).

In the convex case, $(A2')$ is equivalent with the fact that f has bounded and closed level sets, hence admits a minimum point. Suppose that f satisfies $(A2')$ and let $\psi : X \times X \to \mathbb{R}_+$ be as in Subsection 2.2. Consider a bounded sequence (λ_n) of positive real numbers, a sequence (ε_n) of positive real numbers such that, $\sum_{n=0}^{\infty} \varepsilon_n < \infty$, an element $x_0 \in X$ and $e \in \operatorname{int} K$; denote:

$$A_n := \operatorname{argmin}_{x \in X} \left((\varphi_e \circ f)(x) + \lambda_n \psi(x_n, x)\right)$$

and take $x_{n+1} \in X$ s.t. $d(x_{n+1}, A_n) < \varepsilon_n$.

Proposition 3.5. *Suppose, in the above assumptions, that $\lambda_n \to 0$. Then the sequence (x_n) generated as above is bounded and every accumulation point of it is a weak minimizer of f.*

Proof. Under our assumptions the function $\varphi_e \circ f$ is Lipschitz (take into account that φ_e is Lipschitz too) and has weakly compact level sets. We can apply Proposition 2.2 and Lemma 3.2 to get the conclusion. □

We end our discussion by coming back to the infinite dimensional setting and assuming that $\operatorname{int} K = \emptyset$. Then we cannot speak anymore about weak minima, but only about Pareto minima. In this case we make use of another scalarizing functional, namely the function $\tilde{d}_{-K} : Y \cup \{+\infty_K\} \to \mathbb{R} \cup \{+\infty\}$ given by:

$$\tilde{d}_{-K}(y) := \begin{cases} d(y, -K), \text{ if } y \in Y \\ +\infty, \text{ if } y = +\infty_K. \end{cases}$$

It is easy to see that this function is sublinear and Lipschitz on Y. Let $f : X \to Y \cup \{+\infty_K\}$ that satisfies $(A1)$. We have that the composite function $\tilde{d}_{-K} \circ f : X \to \mathbb{R} \cup \{+\infty\}$ is convex and we consider now the IPPA for it: take a bounded sequence of positive real numbers (λ_n), let (ε_n) be a sequence of positive real numbers s.t. $\sum_{n=0}^{\infty} \varepsilon_n < \infty$, consider $x_0 \in X$ arbitrarily, and at every step take $\tilde{x}_{n+1}$ as

$$\tilde{x}_{n+1} = \operatorname{argmin}_{x \in X} \left((\tilde{d}_{-K} \circ f)(x) + \frac{\lambda_n}{2} \|x_n - x\|^2 \right). \tag{3.6}$$

Subsequently consider $x_{n+1} \in X$ s.t. $\|x_{n+1} - \tilde{x}_{n+1}\| < \varepsilon_n$. We have the following result.

Theorem 3.6. *Suppose that f satisfies $(A1)$ and $\tilde{d}_{-K} \circ f$ is lower semicontinuous and has a unique minimum point. Then the sequence generated by (3.6) is well-defined, bounded and weakly convergent to a Pareto minimizer of f.*

Proof. We will prove that the function $\tilde{d}_{-K} \circ f$ is convex. For this, take $y_1, y_2 \in Y$ s.t. $y_1 \leq_K y_2$; this shows that $y_2 - y_1 + K \subset K$ and consequently:

$$\begin{aligned} \tilde{d}_{-K}(y_1) &= \inf_{k \in K} \|y_1 + k\| \leq \inf_{k \in y_2 - y_1 + K} \|y_1 + k\| \\ &= \inf_{k' \in K} \|y_1 + y_2 - y_1 + k'\| = \tilde{d}_{-K}(y_2), \end{aligned}$$

hence $\tilde{d}_{-K}$ is increasing on Y with respect to K. Take now $x_1, x_2 \in X$ and $\alpha \in (0,1)$; if $f(x_1)$ or $f(x_2)$ is $+\infty_K$, there is nothing to prove. Suppose then that $x_1, x_2 \in \operatorname{dom} f$; because of $(A1)$, we have that

$$f(\alpha x_1 + (1-\alpha)x_2) \leq_K \alpha f(x_1) + (1-\alpha) f(x_2),$$

hence $f(\alpha x_1 + (1-\alpha)x_2) \in Y$. Using the fact that $\tilde{d}_{-K}$ is increasing on Y with respect to K and the sublinearity of $\tilde{d}_{-K}$ on Y, we have the convexity of $\tilde{d}_{-K} \circ f$.

Whence, the scalar function $\tilde{d}_{-K} \circ f$ fulfills the conditions in the IPPA, i.e. is convex, lower semicontinuous and admits a minimum. Then (x_n) is well-defined, bounded and weakly convergent to the unique minimum point (say,

$\overline{x}$) of $\tilde{d}_{-K} \circ f$. We show that $\overline{x}$ is a Pareto minimizer of f. Indeed, taking into account the subaditivity of $\tilde{d}_{-K}$ one has that for every $x \in X \setminus \{\overline{x}\}$

$$0 < (\tilde{d}_{-K} \circ f)(x) - (\tilde{d}_{-K} \circ f)(\overline{x}) \leq \tilde{d}_{-K}(f(x) - f(\overline{x}))$$

which yields that $f(x) - f(\overline{x}) \notin -K$ for every $x \in X \setminus \{\overline{x}\}$, whence the conclusion. □

References

[1] H. Bonnel, A. N. Iusem and B. F. Svaiter, *Proximal methods in vector optimization*, SIAM Journal on Optimization, **15** (2005), 953–970.

[2] L.-C. Ceng and J.-C.Yao, *Approximate proximal methods in vector optimization*, European Journal of Operational Research **183** (2007), 1–19.

[3] S. Chrétien and A. O. Hero, *Generalized proximal point algorithms*, Technical Report, University of Michigan, 1998.

[4] M. Durea, *Remarks on strong efficiency in scalar and vector optimization*, Journal of Global Optimization, DOI 10.1007/s10898-009-9453-8.

[5] M. C. Ferris, *Finite termination of the proximal point algorithm*, Mathematical programming, **50** (1991), 359–366.

[6] A. Göpfert, H. Riahi, C. Tammer and C. Zălinescu, *Variational Methods in Partially Ordered Spaces*, Springer, Berlin, 2003.

[7] L. M. Graña Drummond and B. F. Svaiter, *A steepest descent method for vector optimization*, Journal of Computational and Applied Mathematics, **175** (2005), 395–414.

[8] A. N. Iusem, *Augmented Lagrangian methods and proximal point methods for convex optimization*, Investigación Operativa **8** (1999), 11–49.

[9] B. Jiménez and V. Novo, *First and second order sufficient conditions for strict minimality in nonsmooth vector optimization*, Journal of Mathematical Analysis and Applications, **284** (2003), 496–510.

[10] D. T. Luc, *Theory of Vector Optimization*, Springer, Verlag, Berlin, Germany, 1989.

[11] B. Martinet, *Régularisation d'inéquation variationnelles par approximations successives*, Revue Francaise d'Informatique et de Recherche Operationnelle, **3** (1970), 154–179.

[12] R. T. Rockafellar, *Monotone operators and proximal point algorithm*, SIAM Journal on Control and Optimization, **14** (1976), 877–898.

[13] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.