A proximal point method for equilibrium problems in Hilbert spaces

Alfredo N. Iusem* Wilfredo Sosa†

June 25, 2007

Abstract

We propose and analyze a proximal point method for equilibrium problems in Hilbert spaces, which extends the well known proximal point method for variational inequalities. We prove global weak convergence of the generated sequence to a solution of the problem, assuming existence of solutions and rather weak monotonicity properties of the bifunction which defines the equilibrium problem. We also present a reformulation of equilibrium problems as variational inequalities ones, under the same assumptions on f.

Keywords: Equilibrium Problems, Convex Feasibility Problems, Variational Inequalities, Convex Optimization.
Mathematical Subject Classification (2000): 90C47, 49J35.

1 Introduction

Let H be a Hilbert space. Take a closed and convex set $K \subset H$ and $f : K \times K \to \mathbb{R}$ such that

P1: $f(x, x) = 0$ for all $x \in K$,

P2: $f(\cdot, y) : K \to \mathbb{R}$ is upper semicontinuous for all $y \in K$,

P3: $f(x, \cdot) : K \to \mathbb{R}$ is convex and lower semicontinuous for all $x \in K$.

The equilibrium problem $\mathrm{EP}(f, K)$ consists of finding $x^* \in K$ such that $f(x^*, y) \geq 0$ for all $y \in K$. The set of solutions of $\mathrm{EP}(f, K)$ will be denoted as $S(f, K)$.

The equilibrium problem encompasses, among its particular cases, convex optimization problems, variational inequalities (monotone or otherwise), Nash equilibrium problems, and other problems of interest in many applications.

*Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Jardim Botânico, CEP 22460-320 RJ, Rio de Janeiro, Brazil, e-mail: iusp@impa.br

†Universidad Nacional de Ingeniería - Instituto de Matemática y Ciencias Afines, Calle de los Biólogos 245, Lima 12, Lima, Perú, FAX: (51)-1-349-9838, e-mail: sosa@uni.edu.pe

The prototypical example of an equilibrium problem is a variational inequality problem. Since it plays an important role in the sequel, we describe it now in some detail. Consider a continuous $T : H \to H$, and define $f(x,y) = \langle T(x), y-x\rangle$. Then f satisfies P1–P3, and EP(f,K) is equivalent to the variational inequality problem VIP(T,K), consisting of finding a point $x^* \in K$ such that $\langle T(x^*), x - x^*\rangle \geq 0$ for all $x \in K$. We can consider also the case of a point-to-set operator $T : H \to \mathcal{P}(H)$, if it is maximal monotone. In this case VIP(T,K) consists of finding $x^* \in K$ such that that $\langle v^*, x - x^*\rangle \geq 0$ for some $v^* \in T(x^*)$ and all $x \in K$. In this situation, we define $f(x,y) = \sup_{u\in T(x)}\langle u, y-x\rangle$. Though it is less immediate, this f is well defined and it still satisfies P1–P3. Finiteness of f follows from monotonicity of T, and upper semicontinuity of $f(\cdot,y)$ from maximality (via demi-closedness of the graph of maximal monotone operators).

The issue of necessary and/or sufficient conditions for existence of solutions of EP was the starting point in the study of the problem. In 1972, Ky Fan proved existence of solutions assuming compactness of K (see [8]), and a short time afterward the same result was established in [5] assuming instead some form of coerciveness of f.

EP has been extensively studied in recent years, with emphasis on existence results (e.g. [1], [2], [3], [4], [9], [10], [11], [22], [23]). Recently, a new necessary (and in some cases also sufficient) condition for existence of solutions was proposed in [16], and later on simplified and furtherly analyzed in [14]. This condition plays a significant role in our analysis, and appears as condition P5 in Section 2. Its proof is based upon another important theorem by Ky Fan, presented in [7].

Specific algorithms for EP(f,K) do not abound in the literature. Among those of interest, we mention here the methods introduced in [17] and [21]. In this paper, we extend the proximal point method for solving monotone variational inequalities to the case of equilibrium problems in Hilbert spaces. We comment next on this method.

The proximal point algorithm, whose origins can be traced back to [19] and [20], attained its basic formulation in the work of Rockafellar [27]. The algorithm generates a sequence $\{x^k\} \subset H$, starting from some $x^0 \in H$, where x^{k+1} is the unique zero of the operator T^k defined as $T^k(x) = T(x) + \gamma_k(x - x^k)$, with $\{\gamma_k\}$ being a bounded sequence of positive real numbers, called regularization coefficients. It has been proved in [27] that for a maximal monotone T, the sequence $\{x^k\}$ is weakly convergent to a zero of T when T has zeroes, and is unbounded otherwise. Such weak convergence is global, i.e. the result just announced holds in fact for any $x^0 \in H$.

2 Preliminary results

We will need in the sequel certain monotonicity properties of f. We consider the following alternatives:

P4: $f(x,y) + f(y,x) \leq 0$ for all $x, y \in K$.

P4*: Whenever $f(x,y) \geq 0$ with $x, y \in K$, it holds that $f(y,x) \leq 0$.

P4$^\bullet$: There exists $\theta \geq 0$ such that $f(x,y) + f(y,x) \leq \theta \|x-y\|^2$ for all $x, y \in K$.

If $f(x,y) = \sup_{u\in T(x)}\langle u, y-x\rangle$ for some $T : H \to \mathcal{P}(H)$, it is easy to check that P4 is equivalent to monotonicity of T. Thus, a function f satisfying P4 will be said to be *monotone*.

We remind that an operator $T : H \to \mathcal{P}(H)$ is said to be *pseudo-monotone* when $\langle u, x-y\rangle \le 0$ for some $x, y \in H$ and some $u \in T(x)$ implies that $\langle v, x-y\rangle \ge 0$ for all $v \in T(y)$. It is easy to check that if T is pseudo-monotone then f, defined as $f(x,y) = \sup_{u\in T(x)}\langle u, y-x\rangle$, satisfies P4* (the converse statement does not hold in general, but it does when T is point-to-point). For this reason, a function f satisfying P4* will be said to be *pseudo-monotone*. Along the same line, a function f satisfying P4$^\bullet$ will be said to be *θ-undermonotone* (some rationale behind this notation will be presented in Section 4).

We discuss now the relations among P4, P4* and P4$^\bullet$. We start with:

Proposition 1. *Under P1–P3, P4 implies P4*.*

Proof. Elementary. □

On the other hand, P4* does not imply P4, as the following example shows.

Example 1. Let $K = [1/2, 1] \subset \mathbb{R}$ and define $f : K \times K \to \mathbb{R}$ as

$$f(x,y) = x(x-y) \tag{1}$$

Note that $f(x,y) + f(y,x) = (x-y)^2$ so that f is not monotone, but it is immediate that it is 1-undermonotone. The fact that it satisfies P1, P2 and P3 is also immediate. For P4*, note that $f(x,y) \ge 0$ with $x, y \in K$ implies, since $x \ge 1/2$, that $x - y \ge 0$, in which case, using now that $y \ge 1/2$, one has $f(y,x) = y(y-x) \le 0$. This example will be relevant in our analysis.

Now we define our regularization procedure for equilibrium problems. Fix $\gamma \in \mathbb{R}_{++}$ and $\bar{x} \in H$. To any f satisfying P1–P3, we will associate another bifunction $\tilde{f} : K \times K \to \mathbb{R}$ which will be called a *regularization* of f. It is defined as

$$\tilde{f}(x,y) = f(x,y) + \gamma\langle x - \bar{x}, y - x\rangle. \tag{2}$$

The following existence result is essential for establishing that EP$(\tilde{f}, K)$ has solutions.

Proposition 2. *Assume that f satisfies P1, P2, P3, P4*, and additionally the following condition; P5: for any sequence $\{x^n\} \subset K$ satisfying $\lim_{n\to\infty}\|x^n\| = +\infty$, there exists $u \in K$ and $n_0 \in \mathbb{N}$ such that $f(x^n, u) \le 0$ for all $n \ge n_0$;*
then EP(f, K) has solutions.

Proof. See Theorem 4.3 in [14]. □

In view of Proposition 1, the result of Proposition 2 also holds with P4 substituting for P4*. The following two propositions establish some regularizing properties of $\tilde{f}$, as compared to f, under P4$^\bullet$ and P4* respectively. For a convex set $C \subset H$, ri(C) will denote the relative interior of C.

Proposition 3. *Take f satisfying P1, P2, P3 and P4$^\bullet$. Assume that $\gamma > \theta$. Then EP$(\tilde{f}, K)$ has a unique solution.*

Proof. First we prove existence of solutions. We claim that $\tilde{f}$ satisfies the assumptions of Proposition 2. It follows easily from (2) that $\tilde{f}$ inherits P1, P2 and P3 from f. We claim now that $\tilde{f}$ satisfies P4. Note that

$$\tilde{f}(x,y) + \tilde{f}(y,x) = f(x,y) + f(y,x) - \gamma \left\|x-y\right\|^2 \le (\theta - \gamma) \left\|x-y\right\|^2 \le 0, \tag{3}$$

using (2) in the equality, the fact that f satisfies P4$^\bullet$ in the first inequality and the assumption that $\gamma > \theta$ in the second one. In view of Proposition 1, $\tilde{f}$ satisfies P4*. In order to apply Proposition 2, it suffices to establish that $\tilde{f}$ satisfies $P5$. Take a sequence $\{x^n\}$ such that $\lim_{n\to\infty} \|x^n\| = \infty$, and let $u = P_K(\bar{x})$, where $P_K : H \to K$ denotes the orthogonal projection onto K.

Note that

$$\tilde{f}(x^n, u) = f\left(x^n, P_K(\bar{x})\right) + \gamma\langle x^n - \bar{x}, P_K(\bar{x}) - x^n\rangle =$$

$$f\left(x^n, P_K(\bar{x})\right) + \gamma\langle x^n - P_K(\bar{x}), P_K(\bar{x}) - x^n\rangle + \gamma\langle P_K(\bar{x}) - \bar{x}, P_K(\bar{x}) - x^n\rangle \le$$

$$f\left(x^n, P_K(\bar{x})\right) - \gamma \left\|P_K(\bar{x}) - x^n\right\|^2 \le -f\left(P_K(\bar{x}), x^n\right) + \theta \left\|P_K(\bar{x}) - x^n\right\|^2 - \gamma \left\|P_K(\bar{x}) - x^n\right\|^2 =$$

$$-f\left(P_K(\bar{x}), x^n\right) - (\gamma - \theta) \left\|P_K(\bar{x}) - x^n\right\|^2 = -f(u, x^n) - (\gamma - \theta) \left\|u - x^n\right\|^2, \tag{4}$$

using (2) in the first equality, the fact that $\{x^n\} \subset K$, together with the well known "obtuse angle" property of orthogonal projections, in the first inequality, and P4$^\bullet$ in the second inequality. We introduce now some notation for the marginals of f. For each $x \in K$, define $g_x : K \to \mathbb{R}$ as

$$g_x(y) = f(x,y). \tag{5}$$

Take $\hat{x} \in \mathrm{ri}(K)$, so that $\hat{x}$ belongs to the relative interior of the effective domain of g_u. Since g_u is convex by P3, its subdifferential at $\hat{x}$, namely $\partial g_u(\hat{x})$, is nonempty. Take $\hat{v} \in \partial g_u(\hat{x})$. By the definition of subdifferential,

$$\langle \hat{v}, x^n - \hat{x}\rangle \le g_u(x^n) - g_u(\hat{x}) = f(u, x^n) - f(u, \hat{x}). \tag{6}$$

In view of (6),

$$-f(u, x^n) \le \langle \hat{v}, \hat{x} - x^n\rangle - f(u, \hat{x}) \le \|\hat{v}\| \, \|\hat{x} - x^n\| - f(u, \hat{x}) \le$$

$$\|\hat{v}\| \, \|\hat{x} - u\| + \|\hat{v}\| \, \|u - x^n\| - f(u, \hat{x}). \tag{7}$$

Replacing (7) in (4),

$$\tilde{f}(x^n, u) \le \|x^n - u\| \left[\|\hat{v}\| - (\gamma - \theta) \|x^n - u\|\right] + \|\hat{v}\| \, \|\hat{x} - u\| - f(u, \hat{x}). \tag{8}$$

Since $\gamma - \theta > 0$ and $\lim_{n\to\infty} \|x^n\| = \infty$, so that $\lim_{n\to\infty} \|x^n - u\| = \infty$, it follows easily from (8) that $\lim_{n\to\infty} \tilde{f}(x^n, u) = -\infty$, so that $\tilde{f}(x^n, u) \leq 0$ for large enough n. We have verified that $\tilde{f}$ satisfies all the assumptions of Proposition 2, and hence $\mathrm{EP}(\tilde{f}, K)$ has solutions.

Now we prove uniqueness of the solution. Assume that both $\tilde{x}$ and $\tilde{x}'$ solve $\mathrm{EP}(\tilde{f}, K)$. In view of (2),

$$0 \leq \tilde{f}(\tilde{x}, \tilde{x}') = f(\tilde{x}, \tilde{x}') + \gamma\langle \tilde{x} - \bar{x}, \tilde{x}' - \tilde{x}\rangle, \tag{9}$$

$$0 \leq \tilde{f}(\tilde{x}', \tilde{x}) = f(\tilde{x}', \tilde{x}) + \gamma\langle \tilde{x}' - \bar{x}, \tilde{x} - \tilde{x}'\rangle. \tag{10}$$

Adding (9) and (10),

$$0 \leq f(\tilde{x}, \tilde{x}') + f(\tilde{x}', \tilde{x}) - \gamma \left\|\tilde{x} - \tilde{x}'\right\| \leq (\theta - \gamma) \left\|\tilde{x} - \tilde{x}'\right\| \leq 0, \tag{11}$$

using P4 in the second inequality and the fact that $\gamma > \theta$ in the third one. It follows from (11) that $(\theta - \gamma) \|\tilde{x} - \tilde{x}'\| = 0$, and hence $\tilde{x} = \tilde{x}'$, because $\gamma \neq \theta$. □

Proposition 4. *Asume that f satisfies P1, P2, P3 and P4*. If $\tilde{x} \in S(\tilde{f}, K)$ and $x^* \in S(f, K)$ then $\|\tilde{x} - x^*\|^2 + \|\bar{x} - \tilde{x}\|^2 \leq \|\bar{x} - x^*\|^2$*

Proof. Let $\tilde{x}$ be a solution of $\mathrm{EP}(\tilde{f}, K)$ and x^* a solution of $\mathrm{EP}(f, K)$. Since $\tilde{x} \in S(\tilde{f}, K)$, we have

$$0 \leq \tilde{f}(\tilde{x}, x^*) = f(\tilde{x}, x^*) + \gamma\langle \tilde{x} - \bar{x}, x^* - \tilde{x}\rangle,$$

and therefore

$$-f(\tilde{x}, x^*) \leq \gamma\langle \tilde{x} - \bar{x}, x^* - \tilde{x}\rangle, \tag{12}$$

Since x^* solves $\mathrm{EP}(f, K)$, we have $f(x^*, y) \geq 0$ for all $y \in K$. It follows from P4* that $f(y, x^*) \leq 0$ for all $y \in K$, and hence

$$f(\tilde{x}, x^*) \leq 0. \tag{13}$$

Combining (12) and (13),

$$0 \leq \gamma\langle \tilde{x} - \bar{x}, x^* - \tilde{x}\rangle = \frac{\gamma}{2}\left[\|x^* - \bar{x}\|^2 - \|x^* - \tilde{x}\|^2 - \|\tilde{x} - \bar{x}\|^2\right],$$

from which the result follows immediately. □

3 A proximal point method for equilibrium problems

Now we propose the following proximal point method, to be denoted as PPEP, for solving any instance of $\mathrm{EP}(f,K)$ for which f satisfies P1, P2, P3, P4* and P4$^\bullet$. Let θ be the undermonotonicity constant of f. Take a sequence of regularization parameters $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$, for some $\bar{\gamma} > \theta$. Choose $x^0 \in K$ and construct the sequence $\{x^k\} \subset K$ as follows:

Given x^k, x^{k+1} is the unique solution of the problem $\mathrm{EP}(f_k, K)$ where $f_k : K \times K \to \mathbb{R}$ is defined as

$$f_k(x,y) = f(x,y) + \gamma_k \langle x - x^k, y - x \rangle. \tag{14}$$

We mention that if $f(x,y) = \sup_{u \in T(x)} \langle u, y - x \rangle$, for a maximal monotone point-to-set operator $T : H \to \mathcal{P}(H)$, then the sequence defined by (14) is precisely the one generated by the proximal point method for finding solutions of the variational inequality problem $\mathrm{VIP}(T,K)$, studied e.g. in [27].

We present next the convergence result for PPEP. We need first a notion of asymptotic solutions for $\mathrm{EP}(f,K)$. We say that $\{z^k\} \subset K$ is an *asymptotically solving sequence* for $\mathrm{EP}(f,K)$ if $\mathrm{liminf}_{k\to\infty} f(z^k, y) \geq 0$ for all $y \in K$.

Theorem 1. *Consider EP(f,K), where f satisfies P1, P2, P3, P4* and P4$^\bullet$. If EP(f,K) has solutions, then, for all $x^0 \in K$,*

- *i) the sequence $\{x^k\}$ generated by PPEP is bounded and $\lim_{k\to\infty} \left\| x^{k+1} - x^k \right\| = 0$.*
- *ii) $\{x^k\}$ is an asymptotically solving sequence for EP(f,K).*
- *iii) If $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$, then $\{x^k\}$ is weakly convergent to some solution $\hat{x}$ of EP(f,K).*

Proof. i) Since f_k, as defined by (14), is a regularization of f, we obtain, invoking recursively Proposition 3 with $\gamma = \gamma_k > \theta$, $\bar{x} = x^k$ and $\tilde{x} = x^{k+1}$, that the sequence $\{x^k\}$ is well defined (note that for having a well defined sequence we need P4$^\bullet$ but not P4*). Take any $x^* \in S(f,K)$. Since f also satisfies P4*, we invoke Proposition 4 for concluding that

$$\left\| x^{k+1} - x^* \right\|^2 + \left\| x^k - x^{k+1} \right\|^2 \leq \left\| x^k - x^* \right\|^2. \tag{15}$$

It follows that the sequence $\{\left\| x^k - x^* \right\|\}$ is nonnegative and nonincreasing, hence convergent, say to $\sigma \geq 0$. By (15),

$$0 \leq \left\| x^k - x^{k+1} \right\|^2 \leq \left\| x^k - x^* \right\|^2 - \left\| x^{k+1} - x^* \right\|^2. \tag{16}$$

Since the rightmost expression in (16) converges to $\sigma - \sigma = 0$ as $k \to \infty$, we get that

$$\lim_{k\to\infty} (x^k - x^{k+1}) = 0. \tag{17}$$

It is also a consequence of (15) that $\left\|x^k - x^*\right\| \le \left\|x^0 - x^*\right\|$, so that $\{x^k\} \subset B(x^*, \left\|x^0 - x^*\right\|)$, i.e., $\{x^k\}$ is bounded.
ii) Fix any $y \in K$. Since x^{k+1} solves $\mathrm{EP}(f_k, K)$ we have, in view of (14),

$$0 \le f(x^{k+1}, y) + \gamma_k \langle x^{k+1} - x^k, y - x^{k+1}\rangle \le$$

$$f(x^{k+1}, y) + \gamma_k \left\|x^{k+1} - x^k\right\| \left\|y - x^{k+1}\right\|. \tag{18}$$

using Cauchy-Schwartz inequality. We take limits as $k \to \infty$ in (18). Note that $\{\gamma_k\}$ is bounded by $\bar{\gamma}$, $\left\|y - x^{k+1}\right\|$ is bounded by (i), and $\lim_{k\to\infty} \left\|x^{k+1} - x^k\right\| = 0$, also by (i), so that

$$0 \le \mathrm{liminf}_{k\to\infty} f(x^k, y) \ \forall y \in K, \tag{19}$$

and hence $\{x^k\}$ is an asymptotically solving sequence for $\mathrm{EP}(f, K)$, establishing (ii).
iii) In view of (i), $\{x^k\}$ has weak cluster points, all of which belong to K, which, being closed and convex, is weakly closed. Let $\hat{x}$ be one of them. Let $\{x^{j_k}\}$ be a subsequence of $\{x^k\}$ weakly convergent to $\hat{x}$. Under weak upper semicontinuity of $f(\cdot, y)$, we have, in view of (19), $f(\hat{x}, y) \ge \lim_{k\to\infty} f(x^{j_k}, y) \ge 0$ for all $y \in K$, so that $\hat{x} \in S(f, K)$.

It remains to be proved that there exists only one weak cluster point of $\{x^k\}$. Let $\hat{x}$ and $\tilde{x}$ be two weak cluster points of $\{x^k\}$, so that there exist subsequences $\{x^{j_k}\}$ and $\{x^{i_k}\}$ of $\{x^k\}$ whose weak limit points are $\hat{x}$ and $\tilde{x}$ respectively. We have already proved that $\hat{x}$ and $\tilde{x}$ are solutions of $\mathrm{EP}(f, K)$. It follows from (15) that $\{\left\|\hat{x} - x^k\right\|\}$ and $\{\left\|\tilde{x} - x^k\right\|\}$ both converge, say to $\sigma \ge 0$ and $\nu \ge 0$ respectively. Thus

$$2\langle x^{i_k} - x^{j_k}, \tilde{x} - \hat{x}\rangle = \left[\left\|\tilde{x} - x^{i_k}\right\|^2 - \left\|\tilde{x} - x^{j_k}\right\|^2\right] - \left[\left\|\hat{x} - x^{i_k}\right\|^2 - \left\|\hat{x}, x^{j_k}\right\|^2\right] \tag{20}$$

Taking limits as k goes to ∞ in both sides of (20) we get that

$$2\left\|\tilde{x} - \hat{x}\right\|^2 = (\nu - \nu) + (\sigma - \sigma) = 0,$$

and hence $\tilde{x} = \hat{x}$, establishing the uniqueness of the weak accumulation points of $\{x^k\}$. $\square$

At this point, two remarks are in order. Firstly, we comment that the inequality (15) entails the so called *Fejér convergence* of $\{x^k\}$ to the solution set, meaning that the distance from x^k to any solution decreases with k. This property is a consequence of Proposition 4, which is one of the two points in the analysis where assumption P4* is used in an essential way (the other occurs in Proposition 2, where P4* is used for ensuring existence of solutions of the regularized problem).

Also, it deserves to be mentioned that the result of Proposition 2 was proved in [14] under two alternative hypotheses, in addition to P4*, which are called P4 and P4" in this reference. To avoid confusion with the notation of this paper, we will now introduce them as Q4 and Q4'.

Q4: For all $x_1, \ldots, x_n \in K$ and all $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$ it holds that

$$\min_{1 \leq i \leq n} f\left(x_i, \sum_{j=1}^n \lambda_j x_j\right) \leq 0.$$

Q4': For all $x_1, \ldots, x_n \in K$ and $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, it holds that

$$\sum_{i=1}^n \lambda_i f\left(x_i, \sum_{j=1}^n \lambda_j x_j\right) \leq 0.$$

It has been shown in [14] that Q4 in implied by either P4* or Q4', but no additional implication among these three properties hold. It is not hard to prove that Q4 (and henceforth also Q4'), imply that $f(y, x^*) \leq 0$ for all $y \in K$ and all solution x^* of EP(f, K), so that the result of Proposition 4 holds also under Q4 or Q4'. Since, as we have already mentioned, both are also sufficient for that validity of Proposition 2, it happens to be the case that our convergence analysis, and in particular Theorem 1, is still valid if P4* is replaced by either Q4 or Q4'.

Weak upper semicontinuity of $f(\cdot, y)$, as requested in Theorem 1(iii), is quite restrictive, but it holds at least in two significant cases, dealt with in the following corollary.

Corollary 1. *Under the assumptions of Theorem 1,*

i) if H is finite dimensional, then the sequence $\{x^k\}$ generated by PPEP converges to a solution of EP(f, K);

ii) if for all $y \in K$ $f(\cdot, y)$ is concave and can be extended, preserving concavity, to an open set $W \supset K$, then the sequence $\{x^k\}$ generated by PPEP is weakly convergent to a solution of EP(f, K).

Proof. Both results follow from Theorem 1(iii): in the finite dimensional case, weak upper semicontinuity of $f(\cdot, y)$ is just upper continuity, which holds by P2; for (ii), note than concave functions are weakly upper semicontinuous in the relative interior of their effective domain, which turns out to contain K, under the hypothesis of this item. □

In the following section, we will manage to remove the weak upper-semicontinuity assumption, replacing it with a rather weak technical assumption, but only for the monotone (rather than undermonotone) case.

4 A reformulation of the Equilibrium Problem

First we recall our notation for the marginal functions of f given in (5), namely $g_x : K \to \mathbb{R}$ defined, for each $x \in K$, as $g_x(y) = f(x, y)$.

Throughout this section we assume that $\partial g_x(y) \neq \emptyset$ for all $x, y \in K$. This is the case, for instance, if f can be extended, preserving P3, to some open subset V of $H \times H$, containing $K \times K$. We associate to f the operator $T^f : H \to \mathcal{P}(H)$ defined as

$$T^f(x) = \partial g_x(x) + N_K(x), \tag{21}$$

where N_K is the normal operator of K, i.e. the subdifferential of the indicator function I_K, which vanishes at points of K, and takes the value $+\infty$ outside K. The fact that $\partial g_x(x)$ is defined only for $x \in K$ is irrelevant, because $N_K(x) = \emptyset$ when $x \notin K$, and hence the same holds for T^f (one can also think that g_x has been extended to the whole H, taking the value $+\infty$ outside K).

We have the following relation between EP(f, K) and T^f.

Proposition 5. *i) $S(f, K)$ is the set of zeroes of T^f.*

ii) Starting from the same x^0, the sequence generated by PPEP, using f_k as defined by (14), and the sequence generated by the proximal point method for finding zeroes of T^f, coincide (the latter being the sequence $\{x^k\}$, where x^{k+1} is the unique zero of T_k^f, defined as $T_k^f(x) = T^f(x) + \gamma_k(x - x^k)$).

Proof. i) $x^* \in S(f, K)$ iff $g_{x^*}(x^*) = f(x^*, x^*) = 0 \leq f(x^*, y) = g_{x^*}(y)$ for all $y \in K$, i.e. iff x^* solves the problem of minimizing $g_{x^*}(y)$ subject to $y \in K$. The first order condition for this problem, necessary and sufficient for optimality, in view of the convexity of g_{x^*} and K, is the existence of $v^* \in \partial g_{x^*}(x^*)$ such that $\langle v^*, y - x^* \rangle \geq 0$ for all $y \in K$. In view of the definition of N_K, this is precisely equivalent to saying that $0 \in \partial g_{x^*}(x^*) + N_K(x^*)$, i.e., looking at (21), that x^* is a zero of T^f.

ii) Let $\{x^k\}$ be the sequence generated by the proximal point method for finding zeroes of T^f. Assume inductively that x^k is equal to the k-th iterate of PPEP applied to EP(f, K). We must prove that x^{k+1} is the next iterate of the PPEP sequence. We know that

$$0 \in T^f(x^{k+1}) + \gamma_k(x^{k+1} - x^k) = \partial g_{x^{k+1}}(x^{k+1}) + \gamma_k(x^{k+1} - x^k) + N_K(x^{k+1}). \tag{22}$$

For $x \in K$, define $g_x^k : K \to \mathbb{R}$ as

$$g_x^k(y) = g_x(y) + \gamma_k \langle x - x^k, y - x \rangle.$$

It is immediate that $\partial g_x^k(y) = \partial g_x(y) + \gamma_k(x - x^k)$ Define $\tilde{U}^f(y) = \partial g_y^k(y)$. It follows from (22) that x^{k+1} is a zero of $\tilde{U}^f + N_K$, which implies, using now the convexity of g_y^k and of K, that x^{k+1} minimizes $g_{x^{k+1}}^k$ over K, meaning that, for all $y \in K$,

$$0 = g_{x^{k+1}}^k(x^{k+1}) \leq g_{x^{k+1}}^k(y) = g_{x^{k+1}}(y) + \gamma_k \langle x^{k+1} - x^k, y - x^{k+1} \rangle =$$

$$f(x^{k+1}, y) + \gamma_k \langle x^{k+1} - x^k, y - x^{k+1} \rangle = f_k(x^{k+1}, y).$$

Since $0 \leq f_k(x^{k+1}, y)$ for all $y \in K$, x^{k+1} solves $\mathrm{EP}(f_k, K)$, and hence it is the $k+1$-th iterate of the PPEP sequence for problem $\mathrm{EP}(f, K)$, completing the inductive step. □

The result of Proposition 5 can be seen as a converse of our comment at the beginning of Section 2, where we saw that variational inequality problems are particular cases of equilibrium problems: here we have shown that, generally speaking, each equilibrium problem can be reformulated as a variational inequality problem, and, furthermore, that the proximal point method for the equilibrium problem coincides with the classical proximal point method applied to its reformulation as a variational inequality problem. This fact could convey the impression that this whole paper (with the exception, perhaps, of Proposition 5), is rather superfluous, because the proximal point method for variational inequalities, or equivalently for finding zeroes of point-to-set operators, has been extensively analyzed. We argue, however, that such an impression is misleading.

Firstly, convergence results for the classical proximal point method, as presented e.g. in [27], demand monotonicity of the operator, in this case of T^f. Since N_K is always maximal monotone, monotonicity of T^f will occur when the operator $U^f(x) = \partial g_x(x)$ is itself monotone. At this point, it is essential to note that U^f *is not* the subdifferential of a convex function; rather, at each point x it is the subdifferential of a certain convex function, namely g_x, but this function changes with the argument of the operator. Thus, the monotonicity of U is not granted a "priori", but we have the following elementary result.

Proposition 6. *i) If f is θ-undermonotone (i.e. it satisfies P4$^\bullet$) then $U^f + \theta I$ is monotone.*

ii) If f is monotone (i.e. it satisfies P4) then U^f is monotone.

Proof. i) Take $x, y \in K, v \in (U^f + \theta I)(x), w \in (U^f + \theta I)(y)$, so that $v - \theta x \in U^f(x)$, $w - \theta y \in U^f(y)$. Then, using the definition of ∂g_x, (12), P1 and P4$^\bullet$,

$$-\langle (v - \theta x) - (w - \theta y), x - y \rangle = \langle v - \theta x, y - x \rangle + \langle w - \theta y, x - y \rangle \leq$$

$$g_x(y) - g_x(x) + g_y(x) - g_y(y) = f(x, y) - f(x, x) + f(y, x) - f(y, y) =$$

$$f(x, y) + f(y, x) \leq \theta \left\| x - y \right\|^2. \tag{23}$$

It follows easily from (23) that $\langle v - w, x - y \rangle \geq 0$, establishing the monotonicity of $U^f + \theta I$.

ii) Follows from (i) with $\theta = 0$, noting that 0-undermonotonocity is just monotonocity.

□

At this point we emphasize that our convergence analysis in Section 3 holds for instances in which f is not monotone, e.g. the function given in Example 1, and thus the reformulated problem does not fall within the range of the classical proximal point method. We reckon that there are convergence results for the proximal point method applied to non-monotone operators (see e.g. [18]), but still they do not encompass our results here. The closest results seem to be those dealing with hypomonotone operators (e.g. [25], [15], [12]). An operator T is *ρ-hypomonotone* when $T^{-1} + \rho I$ is monotone (I being the identity operator). Now, our assumption of θ-undermonotonicity of f (namely P4$^\bullet$), implies that $U^f + \theta I$ is monotone, but this is different from θ-hypomonotonicity, which means monotonicity of $(U^f)^{-1} + \theta I$. Parenthetically, this could look suspicious, because it is known that monotonicity of $T + \theta I$ is not sufficient for getting convergence of the proximal point method for finding zeroes of T, with the regularization parameters γ_k chosen as $\gamma_k > \theta$, because in general there will be no Fejér convergence of $\{x^k\}$ to the solution set, and (15) will fail. For example, consider EP(f, K) with f as in Example 1, i.e., $f(x, y) = x(x-y)$, but taking now $K = \mathbb{R}$ instead of $K = [1/2, 1]$. Since $f(x, y) + f(y, x) = (x - y)^2$, we have that f is 1-undermonotone also for this choice of K. The only solution of EP(f, K) is now $x^* = 0$, but it is easy to check that if we choose $\gamma_k = \gamma > \theta = 1$, the generated sequence $\{x^k\} \subset \mathbb{R}$ is given by

$$x^k = \left(\frac{\gamma}{\gamma - 1}\right)^k x^0 \tag{24}$$

which diverges for any $x^0 \neq 0$. The point here is that for this choice of K the function f does not satisfy P4*, which ensures the Fejér convergence of $\{x^k\}$ to the solution set.

In the algorithm considered in [15], [25] for finding zeroes of a ρ-hypomonotone operator T, γ_k is required to satisfy $0 \leq \gamma_k < \dfrac{1}{2\rho}$. For f as in Example 1, one has $U^f(x) = -x$, which is 1-hypomonotone, but the analysis in [15], [25] guarantees Fejér convergence of $\{x^k\}$ to $\{0\}$ when $\gamma_k \in (0, 1/2)$, which is inconsistent with the choice prescribed in our algorithm, namely $\gamma_k > \theta = 1$ (note that $\{x^k\}$ as given by (24) indeed converges to 0 with $\gamma \in (0, 1/2)$).

On the other hand, if we take $K = [1/2, 1]$ and $\gamma_k = \gamma > 1$, i.e. satisfying the prescription of the algorithm analyzed in this paper, the divergence effect noted above does not occur, because the sequence is forced to remain in $[1/2, 1]$. Indeed, in this case, due to the presence of the constraint $x^k \in [1/2, 1]$, the iteration formula is not the one given by (24), but rather we have $x^{k+1} = \min\{1, [\gamma/(\gamma - 1)]x^k\}$, and it is easy to check that $\{x^k\}$ converges to the unique solution $x^* = 1$ after a finite number of iterations, for any $x^0 \in K = [1/2, 1]$. The fact that the sequence $\{x^k\}$ does converge to the solution of the problem is consistent with our convergence analysis, since f satisfies P4* for this choice of K.

We mention that we have chosen a one dimensional problem (namely Example 1) for a detailed analysis of the breadth of our results viz a viz others, only because it is posible to get closed iteration formulae, allowing an easy visualization of the sequence behavior. On the other hand, one-dimensional variational inequalities always represent first order conditions of optimization problems. Indeed, Example 1 is equivalent to the first order conditions for the nonconvex problem of minimizing $f(x) = -x^2$ subject to $1/2 \leq x \leq 1$. Convergence results for the proximal point method for nonconvex optimization, appearing in [18], apply to this problem. We give next a two-dimensional example of a pseudomonotone and θ-undermonotone equilibrium problem which is not monotone, and does not represent the first order conditions of an optimization problem.

Example 2. Consider EP(f,K) with $f:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ defined as

$$f(x,y)=(2x_2-x_1)(y_1-x_1),$$

and $K=\{x\in\mathbb{R}^2_+ : \max\{x_1,1/2\}\le x_2\}$, so that K is the quadrangle whose vertices are $(0,1/2),(0,1)$, $(1,1)$ and $(1/2,1/2)$. Note that f is 2-undermonotone, because

$$f(x,y)+f(y,x)=(x_1-y_1)^2-2(x_1-y_1)(x_2-y_2)\le$$

$$(x_1-y_1)^2+\left[(x_1-y_1)^2+(x_2-y_2)^2\right]\le 2\left\|x-y\right\|^2.$$

Also, EP(f,K) is pseudomonotone, because if $f(x,y)\ge 0$, i.e., $(2x_2-x_1)(y_1-x_1)\ge 0$, then

$$y_1-x_1\ge 0, \tag{25}$$

because for all $x\in K$ one has $x_2\ge x_1>0$, which gives $2x_2-x_1>0$. In such a case, since $2y_2-y_1>0$ for all $y\in K$, by the same argument as above, and $x_1-y_1\le 0$ by (25), one gets $f(y,x)=(2y_2-y_1)(x_1-y_1)\le 0$, establishing pseudomonotocity. The choice $x=(1,1)$, $y=(0,1)$, gives $f(x,y)+f(y,x)=1$ with $x,y\in K$, so that the problem is not monotone. Finally, it does not fall within the optimization case, because its associated operator U^f is given by $U^f(x)=Ax$ with

$$A=\begin{vmatrix}-1 & 2\\ 0 & 0\end{vmatrix},$$

which is not symmetric. In such a case, the variational inequality problem VIP(U^f,K), equivalent to the problem of finding zeroes of $T^f=U^f+N_K$, cannot represent the first order optimality conditions of any optimization problem.

In summary, our analysis works for $\gamma_k\ge\theta$, where θ is the undermonotonicity constant, because P4* comes to the rescue: it allows us to prove Proposition 4, for which P4$^\bullet$ is not enough. To our knowledge, the convergence of the proximal point method for a pseudo-monotone operator T which is also θ-undermonotone (in the sense that $T+\theta I$ is maximal monotone) with $\gamma_k\ge\theta$, has not been studied up to now. In fact, such a convergence analysis follows, up to certain technicalities, from the results in this paper, but we will not pursue this issue further. We just mention that a one-dimensional example of a pseudo-monotone and θ-undermonotone operator which is not monotone is given by $T:\mathbb{R}\to\mathcal{P}(\mathbb{R})$ defined as $T(x)=-\theta x+N_{[\alpha,\beta]}(x)$ with $0<\alpha<\beta$, and a two-dimensional one, with $\theta=2$, is given by the operator T^f associated to Example 2. The end of the story is that the current literature on the proximal point method does not cover a case like the one given in Example 2.

We give also at this point a reason for denoting as θ-undermonotone, and not θ-hypomonotone, a function f satisfying P4$^\bullet$: this property translates into monotonicity of the operator $U^f+\theta I$, which, as explained above, does not coincide with θ-hypomonotocity of U^f, as defined above, following e.g., [25].

In addition, not only monotonicity of the operator (or some variant thereof) is needed in the analysis of the proximal point method, but also maximality. Even in the case in which f is monotone and defined on the whole space H, it is not at all obvious that the operator U^f will be maximal monotone. We remind that while monotonicity of the subdifferential of a convex function is immediate, its maximality is rather nontrivial (see [26]). We will prove below that the needed maximality indeed holds when f is monotone, but it happens that the proof requires the EP techniques: it uses in an essential way the existence result in Proposition 2. Since we cannot avoid the Equilibrium Problem approach even when dealing directly with the reformulation, we considered it advisable to present a clean analysis of the PPEP, just in terms of the Equilibrium Problem, as done in Section 3, going as far as possible without introducing the more complicated machinery of the reformulation.

However, we proceed now to explore the reformulation in order to remove the weak upper-semicontinuity hypothesis.

We present next a result on maximality of monotone operators, of some interest on its own. A similar result, in reflexive Banach spaces but with $\lambda = 1$, was established in Remark 10.8 of [28]. We present here a simplified version of the proof of Theorem 4.5.7 in [6], which also deals with Banach spaces.

Proposition 7. *Let $T : H \to \mathcal{P}(H)$ be a monotone operator. If $T + \lambda I$ is onto for some $\lambda > 0$, then T is maximal monotone.*

Proof. Take a monotone operator $\bar{T}$ such that $T \subset \bar{T}$, and a pair (v, z) such that $v \in \bar{T}(z)$. We must prove that $v \in T(z)$. Define $b = v + \lambda z$. Since $T + \lambda I$ is onto, there exists $x \in H$ such that

$$b \in T(x) + \lambda x \subset \bar{T}(x) + \lambda x. \tag{26}$$

On the other hand, since $v \in \bar{T}(z)$, we have that $b = v + \lambda z \in \bar{T}(z) + \lambda z$. Since $\bar{T} + \lambda I$ is strictly monotone, we conclude that $x = z$, and thus, making $x = z$ in the first inclusion of (26), we have $v + \lambda z \in T(z) + \lambda z$, which implies that $v \in T(z)$. It follows that $\bar{T} \subset T$, i.e. $\bar{T} = T$, and hence T is maximal. ☐

Now we use Propositions 2 and 7 to establish maximal monotonicity of T^f under adequate assumptions on f.

Proposition 8. *If f satisfies P1–P4, then T^f, as defined by* (21)*, is maximal monotone.*

Proof. We intend to apply Proposition 7, for which we need to show that T^f is monotone and that $T^f + \lambda I$ is onto for some $\lambda > 0$. Note that U^f, defined as $\tilde{U}^f(y) = \partial g_y^k(y)$, is monotone by Proposition 6(i) with $\theta = 0$ (recall that 0-undermonotonicity is just monotonicity). Since N_K is certainly monotone, it follows that $T^f = U^f + N_K$ is monotone. Now we address the surjectivity issue. Take any $\lambda > 0$ and $b \in H$. We want to show that there exists $x \in K$ such that $b \in (T^f + \lambda I)x$. Consider $\tilde{f}$ as in (2), with $\bar{x} = \lambda^{-1} b$, $\gamma = \lambda$. By Proposition 3, EP$(\tilde{f}, K)$ has a solution, say x. Define

$$\tilde{g}_x(y) = \tilde{f}(x, y) = f(x, y) + \lambda \langle x - \lambda^{-1} b, y - x \rangle. \tag{27}$$

Note that, since x solves EP$(\tilde{f}, K)$,

$$\tilde{g}_x(x) = \tilde{f}(x, x) = 0 \leq \tilde{f}(x, y) = \tilde{g}_x(y)$$

for all $y \in K$. Thus, x minimizes $\tilde{g}_x$ over K, which is the same as saying that x is an unrestricted minimizer of $g_x + I_K$, where I_K is the indicator function of K. By assumption, we have that $\partial g_x(z)$, and henceforth $\partial \tilde{g}_x(z)$, are nonempty for all $z \in K$. In view of (27) and the fact that $\partial I_K = N_K$, we have

$$0 \in \partial(\tilde{g}_x + I_K)(x) = \partial g_x(x) + \lambda(x - \lambda^{-1} b) + N_K(x) = \partial g_x(x) + \lambda x - b + N_K(x). \tag{28}$$

Rewriting (28) as

$$b \in \partial g_x(x) + N_K(x) + \lambda x = (U^f + N_K)(x) + \lambda x = (T^f + \lambda I)(x),$$

we complete the proof of surjectivity of $T^f + \lambda I$. We can use now Proposition 7 to conclude that T^f is maximal monotone. □

We remark that in the case of $K = H$, and $f(x, y) = h(y) - h(x)$, where $h : H \to \mathbb{R} \cup \{+\infty\}$ is a convex function, we get $\partial g_x = \partial h$ and $N_K(x) = 0$ for all $x \in H$, so that $T^f = U^f = \partial h$. Since this f satisfies P1–P4, Proposition 8 provides an alternative (and rather short) proof of the maximality of the subdifferential of a convex function (cf. [26]). Noting that the proof of Proposition 7 is also quite short, we conclude that the "heavy artillery" behind this approach is hidden in the proof of Proposition 2, given in [14].

We remind now a well known property of maximal monotone operators, namely demi-closedness.

Definition 1. *Given $T : H \to \mathcal{P}(H)$, the graph of T is said to be demi-closed, when the following property holds: if $\{x^k\} \subset H$ is weakly convergent to $x^* \in H$, $\{v^k\} \subset H$ is strongly convergent to $v^* \in H$, and $v^k \in T(x^k)$ for all k, then $v^* \in T(x^*)$.*

Proposition 9. *If $T : H \to \mathcal{P}(H)$ is maximal monotone, then its graph is demi-closed.*

Proof. See, e.g., [24], p. 105. □

Now we can get rid of the weak upper semi-continuity assumption in Theorem 1(iii).

Theorem 2. *If f satisfies P1–P4, EP(f, K) has solutions, and $f(x, \cdot)$ can be extended, for all $x \in K$, to an open set $W \supset K$, while preserving its convexity, then the sequence $\{x^k\}$ generated by PPEP is weakly convergent to a solution of EP(f, K) for all $x^0 \in K$.*

Proof. Note that we are within the assumptions of items (i) and (ii) of Theorem 1, recalling that P4 implies $P4^*$. Define $g_x^k(y) = f_k(x, y)$, with f_k as in (14). As we have already seen several times, x^{k+1} is a solution of the problem min $g^k_{x^{k+1}}(y)$ subject to $y \in K$, and thus it satisfies the first order optimality condition, namely

$$0 \in \partial g^k_{x^{k+1}}(x^{k+1}) + N_K(x^{k+1}) = \partial g_{x^{k+1}}(x^{k+1}) + \gamma_k(x^{k+1} - x^k) + N_K(x^{k+1}),$$

which can be rewritten as

$$v^{k+1} := \gamma_k(x^k - x^{k+1}) \in \partial g_{x^{k+1}}(x^{k+1}) + N_K(x^{k+1}) = T^f(x^{k+1}). \tag{29}$$

Note that T^f is maximal monotone by Proposition 8, so that its graph is demiclosed by Proposition 9. Also, $\{x^k\}$ is bounded by Theorem 1(i). Observe also that $\{v^k\}$ is strongly convergent to 0 by Theorem 1(i) and boundedness of $\{\gamma_k\}$. Let $\hat{x}$ be a weak cluster point of $\{x^k\}$. Taking limits along the corresponding subsequence in (29), we are exactly in the situation of Definition 1, so that Proposition 9 entails that $0 \in T^f(\hat{x})$. By Proposition 5(i), $\hat{x}$ solves EP(f, K). Uniqueness of the cluster points of $\{x^k\}$, and consequently weak convergence of $\{x^k\}$ to a point in $S(f, K)$, follow with the argument used at the end of the proof of Theorem 1(iii). □

We remark that the difference between Theorems 1 and 2, besides the fact that the proof of the latter requires the reformulation of the equilibrium problem as a variational inequality one, lies in the assumptions on f (monotonicity and extension of $f(x, \cdot)$ to an open set containing K), which replace weak upper semicontinuity of $f(\cdot, y)$, as the tool for establishing optimality of the weak cluster points of the generated sequence.

At this point it would be reasonable to discuss the convergence of the method under inexact solution of the subproblems, which is the standard situation in actual implementations. Different error criteria which preserve the convergence result for the proximal point method for finding zeroes of maximal monotone operators have been proposed since Rockefellar's 1976 paper ([27]). Recently, less stringent error criteria, allowing for constant relative errors along the iterations, were proposed by Solodov and Svaiter in [29], [30] for monotone operators, and extended to hypomonotone operators in [15] (in the case of Hilbert spaces) and [12] (in the case of Banach spaces). Through the reformulation of equilibrium problems as variational inequality problems, proposed in this section, all these convergence results hold for equilibrium problems, assuming, of course, that the monotonicity properties of f are such that the associated operator T^f satisfies the assumptions required for the convergence of each of these inexact procedures. Nevertheless, it is possible to introduce some error criteria which are specific of equilibrium problems. These criteria will be the subject of a forthcoming paper.

References

[1] Bianchi, M., Pini, R. A note on equilibrium problems with properly quasimonotone bifunctions. *Journal of Global Optimization* **20** (2001) 67-76.

[2] Bianchi, M., Pini, R. Coercivity conditions for equilibrium problems. *Journal of Optimization Theory and Applications* **124** (2005) 79-92.

[3] Bianchi, M., Schaible, S. Generalized monotone bifunctions and equilibrium problems. *Journal of Optimization Theory and Applications* **90** (1996) 31-43.

[4] Blum, E. Oettli, W. From optimization and variational inequalities to equilibrium problems. *The Mathematics Student* **63** (1994) 123-145.

[5] Brezis, H., Nirenberg, L., Stampacchia, S. A remark on Ky Fan minimax principle. *Bolletino della Unione Matematica Italiana* **6** (1972) 293-300.

[6] Burachik, R.S., Iusem, A.N. *Set-Valued Mappings and Enlargements of Monotone Operators*. Springer, Berlin (2007).

[7] Fan, K. A generalization of Tychonoff's fixed point theorem. *Mathematische Annalen* **142** (1961) 305-310.

[8] Fan, K. A minimax inequality and applications. In *Inequality III* (O. Shisha, editor) Academic Press, New York (1972) 103-113.

[9] Flores-Bazán, F. Existence theorems for generalized noncoercive equilibrium problems: quasi-convex case, *SIAM Journal on Optimization* **11** (2000) 675-790.

[10] Flores-Bazán, F. Ideal, weakly efficient solutions for convex vector optimization problems. *Mathematical Programming* **93** (2002) 453-475.

[11] Flores-Bazán, F. Existence theory for finite dimensional pseudomonotone equilibrium problems. *Acta Applicandae Mathematicae* **77** (2003) 249-297.

[12] Gárciga Otero, R., Iusem, A.N. Proximal methods in Banach spaces without monotonicity. *Journal of Mathematical Analysis and Applications* **330** (2007) 433-450.

[13] Hadjisavvas, N. Schaible, S. On strong pseudomonotonicity and (semi)strict quasimonotonicity. *Journal of Optimization Theory and Applications* **79** (1993) 139-155.

[14] Iusem, A.N., Kassay, G., Sosa, W. On certain conditions for the existence of solutions of equilibrium problems (to be published in *Mathematical Programming*).

[15] Iusem, A.N., Pennanen, T., Svaiter, B.F. Inexact variants of the proximal point method without monotonicity. *SIAM Journal on Optimization* **13** (2003) 1080-109.

[16] Iusem, A.N., Sosa, W. New existence results for equilibrium problems. *Nonlinear Analysis* **52** (2003) 621-635.

[17] Iusem, A.N., Sosa, W. Iterative algorithms for equilibrium problems. *Optimization* **52** (2003) 301-316.

[18] Kaplan, A., Tichatschke, R. Proximal point methods and nonconvex optimization. *Journal of Global Optimization* **13** (1998) 389-406.

[19] Krasnoselskii, M.A. Two observations about the method of succesive approximations. *Uspekhi Matematicheskikh Nauk* **10** (1955) 123-127.

[20] Moreau, J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Societé Mathématique de France* **93** (1965).

[21] Muu, L.D., Oettli, W. Convergence of an adaptive penalty scheme for finding constraint equilibria. *Nonlinear Analysis* **18** (1992) 1159-1166.

[22] Oettli, W. A remark on vector-valued equilibria and generalized monotonicity. *Acta Mathematica Vietnamita* **22** (1997) 215-221.

[23] Oettli, W., Schläger, D. Generalized vectorial equilibrium and generalized monotonicity. In *Functional Analysis with Current Applications in Science, Technology and Industry (Aligarh, 1996). Pitman Research Notes in Mathematics Series* **377** (1998) 145-154.

[24] Pascali, D., Sburlan, S. *Nonlinear Mappings of Monotone Type*. Editura Academiei, Bucarest (1978).

[25] Pennanen, T. Local convergence of the proximal point method and multiplier methods without monotonicity. *Mathematics of Operations Research* **27** (2002) 170-191.

[26] Rockafellar, R.T. On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics* **33** (1970) 209-216.

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

[28] Simons, S. *Minimax and Monotonicity*. Lecture Notes in Mathematics **1693**. Springer, Berlin (1998).

[29] Solodov, M.V., Svaiter, B.F. A hybrid projection–proximal point algorithm. *Journal of Convex Analysis* **6** (1999) 59-70.

[30] Solodov, M.V., Svaiter, B.F. An inexact hybrid extragardient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis* **7** (1999) 323-345.