Journal of Convex Analysis
Volume 18 (2011), No. 2, 489–503

On the Maximal Monotonicity of Diagonal Subdifferential Operators

A. N. Iusem*
Instituto de Matemática Pura e Aplicada (IMPA),
Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil
iusp@impa.br

Received: July 8, 2009

Consider a real-valued bifunction f which is concave in its first argument and convex in its second one. We study its subdifferential with respect to the second argument, evaluated at pairs of the form (x, x), and the subdifferential of $-f$ with respect to its first argument, evaluated at the same pairs. The resulting operators are not always monotone, and we analyze additional conditions on f which ensure their monotonicity, and furthermore their maximal monotonicity. Our main result is that these operators are maximal monotone when f is continuous and it vanishes whenever both arguments coincide. Our results have consequences in terms of the reformulation of equilibrium problems as variational inequality ones.

Keywords: Equilibrium problem, maximal monotone operator, diagonal subdifferential

2000 Mathematics Subject Classification: 90C47, 49J35

1. Introduction

Let X be a reflexive Banach space and X^* its dual. Consider a function $f : X \times X \to \mathbb{R}$ which is concave in its first argument and convex in its second one. Let $\partial_1(-f)(x, y), \partial_2 f(x, y)$ denote the subdifferentials of $-f$ with respect to its first argument, and of f with respect to the second one, respectively, evaluated at a point $(x, y) \in X \times X$. The well known *saddle point operator* $T_f : X \times X \to \mathcal{P}(X^* \times X^*)$ is defined as

$$T_f(x, y) = (\partial_1(-f)(x, y), \partial_2 f(x, y)). \tag{1}$$

We will be concerned in this paper with two other set-valued operators related to the bifunction f, namely $R_f, S_f : X \to \mathcal{P}(X^*)$, defined as:

$$R_f(x) = \partial_1(-f)(x, x), \tag{2}$$

$$S_f(x) = \partial_2 f(x, x). \tag{3}$$

We will refer to R_f, S_f as *diagonal subdifferential* operators. Obeserve that neither R_f nor S_f are subdifferentials of convex functions: at each point x each one of them coincides with the subdifferential of a certain convex function evaluated at x, but the

*The work of this author was partially supported by CNPq grant no. 301280/86.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

functions themselves change with x. More precisely, $S_f(x)$ is the subdifferential of the convex function $f_x : X \to \mathbb{R}$ evaluated at x, where f_x is defined as $f_x(y) = f(x, y)$. Similarly, $R_f(x)$ is the subdifferential of the convex function $-f(\cdot, x)$ evaluated at x. In fact, as we will show later on, both R_f and S_f may fail to be monotone operators, unless additional assumptions are imposed upon f. The study of these conditions is the purpose of this paper.

The motivation for studying these operators arises from the so called *equilibrium problem*, which we describe next. Given X, f as above (possibly with additional and/or slightly different assumptions on f, some of which will be detailed later on), and a closed and convex subset $C \subset X$, the equilibrium problem $\mathrm{EP}(f, C)$ consists of finding $x^* \in C$ such that $f(x^*, x) \geq 0$ for all $x \in C$. See [2], [8] and [7] for definitions and properties of equilibrium problems pertinent to the subject of this paper.

Under the additional assumption that $f(x, x) = 0$ for all $x \in X$, the convexity of $f(x, \cdot)$ implies easily that x^* solves $\mathrm{EP}(f, C)$ if and only x^* minimizes the marginal function f_{x^*} defined above on the feasible set C, which happens if and only if x^* is a zero of the sum of the subdifferential of this objective funcion and the normalized cone N_C of C, i.e. a zero of $S_f + N_C$. Equivalently, x^* is a solution of the variational inequality problem $\mathrm{VIP}(S_f, C)$. It is well known that variational inequality problems are substantially easier to solve when the involved operator is maximal monotone. Thus, the study of conditions under which S_f is maximal monotone has a significant impact on the theory of equilibrium problems. We remind here that a set-valued operator $T : X \to \mathcal{P}(X^*)$ is *monotone* if $\langle u_1 - u_2, x_1 - x_2 \rangle \geq 0$ for all $(x_1, u_1), (x_2, u_2) \in G(T)$, where the *graph* $G(T)$ of T is defined as $G(T) = \{(x, u) \in X \times X^* : u \in T(x)\}$. T is said to be *maximal monotone* if it is monotone and $G(T) = G(T')$ for all monotone operator $T' : X \to \mathcal{P}(X^*)$ such that $G(T) \subset G(T')$.

We will prove in this paper that R_f and S_f are monotone under some further assumptions on the bifunction f, besides its concave-convex property, related to its behavior as a funtion of its two arguments simultaneously, like for instance being jointly continuous on x, y and vanishing on the diagonal of $X \times X$. Monotonicity of S_f will also be established without demanding concavity of $f(\cdot, y)$, but imposing instead stronger joint assumptions on f: it must vanish on the diagonal and be a *monotone bifunction*, meaning that $f(x, y) + f(y, x) \leq 0$ for all $(x, y) \in X \times X^*$. Similarly, it will be proved that R_f is monotone when $-f$ is monotone and it vanishes on the diagonal, without requiring convexity of $f(x, \cdot)$. These results will be proved in Section 2.

In Section 3 we deal with the maximal monotonicity of S_f and R_f. We prove that both operators are indeed maximal monotone under any of our sets of assumptions guaranteeing their monotonicity, with an additional hypothesis on the Banach space X. We describe next this result and the main idea behind its proof. We recall that the *duality operator* $J : X \to \mathcal{P}(X^*)$ is the subdifferential of the convex function $(1/2) \|\cdot\|^2$. A well known result, due to R. T. Rockafellar (see Theorem 4.4.7 in [4]), states that, in a Banach space X such that both J and J^{-1} are single-valued, a monotone operator T is maximal monotone if and only if the operator $T + J$ is surjective. Thus, we will try to prove surjectivity of $S_f + J$ (same for R_f). For this surjectivity result, we will exhibit some $\tilde{f}$ such that $S_{\tilde{f}} = S_f + J$, and then apply a theorem on existence of solutions of equilibrium problems, related to some results proved in [8], [7], to the problem EP$(\tilde{f}, X)$, which easily implies surjectivity of $S_f + J$. The appropriate $\tilde{f}$ is given by

$$\tilde{f}(x,y) = f(x,y) + \frac{1}{2}\left(\|y\|^2 - \|x\|^2\right) + \langle b, x-y\rangle,$$

where $\langle\cdot,\cdot\rangle$ denotes the duality coupling in $X^* \times X$, and b is a fixed element of X^*. It turns out to be the case that $\tilde{f}$ inherits from f all the assumptions used in our analysis, like concavity-convexity, monotonicity, continuity, etc.

We close this section with some comments on the operator T_f defined by (1). It is easy to check that when the bifunction f is concave-convex then the zeroes of T_f are the saddle points of f, i.e. $(0,0) \in T_f(x^*, y^*)$ if and only if

$$f(x, y^*) \leq f(x^*, y^*) \leq f(x^*, y)$$

for all $(x,y) \in X \times X$.

Also, we mention that the most important example of a concave-convex bifunction is the Lagrangian function L associated to the convex minimization problem,

$$\min h_0(y) \qquad \text{s.t.} \;\; h_i(y) \leq 0 \;\; (1 \leq i \leq m),$$

with $h_i : \mathbb{R}^n \to \mathbb{R}$ convex $(0 \leq i \leq m)$. In this case, $L : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is defined as $L(x,y) = h_0(y) + \sum_{i=1}^m x_i h_i(y)$. L is clearly concave in x for all $y \in \mathbb{R}^n$, and convex in y for all $x \in \mathbb{R}^n_+$. The solutions of VIP$(T_L, \mathbb{R}^n \times \mathbb{R}^m_+)$ are optimal primal-dual pairs for the optimization problem above.

2. Monotonicity of the diagonal subdifferential operators

We start by introducing the concave-convex property of the bifunction f in a formal way. We consider the following two assumptions on $f : X \times X \to \mathbb{R}$.

A1) $f(\cdot, y)$ is concave for all $y \in X$.
A2) $f(x, \cdot)$ is convex for all $x \in X$.

We will prove first that, assuming only A1) and A2), the operator $R_f + S_f$ is monotone. This result will follow from the monotonicity of T_f under the same assumptions. It is well known that T_f is monotone, and furthermore maximal monotone (see e.g. Theorem 4.7.5 in [4]), but we include a proof of this fact, which is quite elementary, for the sake of self-containment, and also because we will use part of it later on.

Proposition 2.1. *Assume that X is a reflexive Banach space and that $f : X \times X \to \mathbb{R}$ satisfies* A1) *and* A2) *above. Consider T_f, S_f and R_f as defined by* (1), (2) *and* (3) *respectively. Then*

i) T_f *is monotone.*
ii) $R_f + S_f$ *is monotone.*

Proof. *i)* Take $(x_i, y_i) \in X \times X$, $(u_i, v_i) \in T(x_i, y_i)$ $(i = 1, 2)$. We must verify that

$$\langle (u_1, v_1) - (u_2, v_2), (x_1, y_1) - (x_2, y_2)\rangle \geq 0. \tag{4}$$

Note that $u_i \in \partial_1(-f)(x_i, y_i), v_i \in \partial_2 f(x_i, y_i)$ $(i = 1, 2)$. By convexity of $-f(\cdot, y_i)$:

$$\langle -u_1, x_1 - x_2\rangle = \langle u_1, x_2 - x_1\rangle \le -f(x_2, y_1) + f(x_1, y_1), \tag{5}$$

$$\langle u_2, x_1 - x_2\rangle \le -f(x_1, y_2) + f(x_2, y_2). \tag{6}$$

Adding (5) and (6):

$$-\langle u_1 - u_2, x_1 - x_2\rangle \le f(x_1, y_1) - f(x_1, y_2) - f(x_2, y_1) + f(x_2, y_2). \tag{7}$$

By convexity of $f(x_i, \cdot)$:

$$\langle -v_1, y_1 - y_2\rangle = \langle v_1, y_2 - y_1\rangle \le f(x_1, y_2) - f(x_1, y_1), \tag{8}$$

$$\langle v_2, y_1 - y_2\rangle \le f(x_2, y_1) - f(x_2, y_2). \tag{9}$$

Adding (8) and (9):

$$-\langle v_1 - v_2, y_1 - y_2\rangle \le -f(x_1, y_1) + f(x_1, y_2) + f(x_2, y_1) - f(x_2, y_2). \tag{10}$$

Adding (7) and (10), and multiplying by -1,

$$0 \le \langle u_1 - u_2, x_1 - x_2\rangle + \langle v_1 - v_2, y_1 - y_2\rangle = \langle (u_1, v_1) - (u_2, v_2), (x_1, y_1) - (x_2, y_2)\rangle, \tag{11}$$

establishing (4) and the monotonicity of T_f.

ii) Take $z_i \in (R_f + S_f)(x_i)$ $(i = 1, 2)$. Then $z_i = u_i + v_i$ with $u_i \in R_f(x_i), v_i \in S_f(x_i)$, i.e. $u_i \in \partial_1(-f)(x_i, x_i), v_i \in \partial_2 f(x_i, x_i)$. It follows that $z_i = (u_i, v_i) \in T_f(x_i, x_i)$ $(i = 1, 2)$ and therefore,

$$\langle z_1 - z_2, x_1 - x_2\rangle = \langle u_1 - u_2, x_1 - x_2\rangle + \langle v_1 - v_2, x_1 - x_2\rangle \ge 0, \tag{12}$$

using (11) with $x_i = y_i$ $(i = 1, 2)$. In view of (12), $R_f + S_f$ is monotone. □

We remark now that under just A1) and A2), the operators R_f, S_f may fail to be monotone. Take $X = \mathbb{R}^n$, and an indefinite $A \in \mathbb{R}^{n\times n}$, i.e. such that there exist $\tilde{x}, \hat{x} \in \mathbb{R}^n$ satisfying $\tilde{x}^t A\tilde{x} > 0$, $\hat{x}^t A\hat{x} < 0$. Define $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ as $f(x, y) = x^t Ay$, so that $T_f(x, y) = (-Ay, A^t x), R_f(x) = -Ax, S_f(x) = A^t x$, and hence $(R_f + S_f)(x) = (-A + A^t)(x)$. Since $-A + A^t$ is skew-symmetric, $R_f + S_f$ is indeed monotone, as established in Proposition 2.1 *ii*), but the indefiniteness of A implies that neither R_f nor S_f are monotone.

We introduce next additional assumptions on the joint behavior of f in its two arguments, which will allow us to establish monotonicity of R_f, S_f. The key property seems to be the following: f must be constant on the diagonal of $X \times X$, i.e., there must exist $\mu \in \mathbb{R}$ such that $f(x, x) = \mu$ for all $x \in X$. Since both R_f and S_f are defined up to additive constants in f, without loss of generality we will assume that $\mu = 0$.

Something else is needed, and at this point we will consider two alternatives. The first of them consists of demanding monotonicity of either f or $-f$. We recall that a bifunction $f : X \times X \to \mathbb{R}$ is said to be *monotone* if

$$f(x, y) + f(y, x) \le 0 \tag{13}$$

for all $(x, y) \in X \times X$.

We will consider the following assumptions related to monotoncity of f.

A3) $f(x,x) = 0$ for all $x \in X$.
A4) f is monotone.
A5) $-f$ is monotone.

Working under these assumptions, we can relax the concavity-convexity hypotheses on f: we will need only convexity of $f(x,\cdot)$, i.e. A2), for proving monotonicity of S_f, and just concavity of $-f(\cdot,y)$, i.e. A1), for monotonicity of R_f.

A second and more interesting alternative consists of avoiding any monotonicity assumption on f, and instead adding to the concavity-convexity properties given by A1), A2), a rather weak assumption on the joint behavior of f in its two arguments, namely

A6) f is continuous on $X \times X$.

We will prove monotonicity of R_f under A1), A3) and A5), and of S_f under A2), A3) and A4), along the first alternative, and later on we will follow the second alternative, establishing monotonicity of both R_f and S_f under A1), A2), A3) and A6).

We mention that none of these two sets of assumptions implies the remaining one. We give two examples, both of them with $X = \mathbb{R}^n$. Take $A, C \in \mathbb{R}^{n\times n}$ positive semidefinite, but such that $A - C$ is indefinite. Define

$$f(x,y) = -x^t Ax + y^t Cy + x^t(A-C)y. \tag{14}$$

This f satisfies A1), A2), A3) and A6), but not A4), because neither f nor $-f$ is monotone: note that $f(x,y)+f(y,x) = (x-y)^t(C-A)(x-y)$, which is neither positive nor negative for all $x, y \in X$, due to the indefiniteness of $A - C$. A non-quadratic example with the same properties is obtained by taking $\bar{f}(x,y) = f(x,y)-h(x)+h(y)$, with f as in (14), where $h : X \to \mathbb{R}$ is an arbitrary convex function.

Consider now $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x,y) = \sum_{j=1}^{n} x_j^3(y_j - x_j).$$

This f satisfies A2), A3) and A4) (note that $f(x,y) + f(y,x) = \sum_{j=1}^{n}(x_j^3 - y_j^3)(y_j - x_j) \le 0$), but A1) fails, because f is not concave in x for all y. The bifunction $-f$, with f as in this example, satisfies A1), A3) and A5), but not A2).

At this point it is convenient to formalize a certain symmetry relation between R_f and S_f. To any bifunction $f : X \times X \to \mathbb{R}$ we associate the bifunction $g : X \times X \to \mathbb{R}$ defined as $g(x,y) = -f(y,x)$. The connections between R_f, S_f, R_g and S_g are encapsulated in the following proposition.

Proposition 2.2.

i) f *satisfies* A1) *iff* g *satisfies* A2),
ii) f *satisfies* A2) *iff* g *satisfies* A1),
iii) f *satisfies* A3) *iff* g *satisfies* A3),
iv) f *satisfies* A4) *iff* g *satisfies* A5),
v) f *satisfies* A5) *iff* g *satisfies* A4),

vi) $R_f = S_g$, $S_f = R_g$.

Proof. Elementary. □

Proposition 2.2 will allow us to obtain results for R_f from similar results for S_f, avoiding the duplication of arguments.

We have the following results on monotonicity of R_f, S_f, assuming monotonicity properties of f.

Theorem 2.3.

i) *If f satisfies* A1), A3) *and* A5) *then R_f is monotone.*
ii) *If f satisfies* A2), A3) *and* A4) *then S_f is monotone.*

Proof. *i)* Take $x_i \in X$, $u_i \in R_f(x_i)$ $(i = 1, 2)$. In view of A1), which is equivalent to convexity of $-f(\cdot, x_i)$, we invoke (7), obtained under just this assumption, with $x_i = y_i$, and we get

$$\langle u_1 - u_2, x_1 - x_2 \rangle \geq -f(x_1, x_1) + f(x_1, x_2) + f(x_2, x_1) - f(x_2, x_2) \geq 0, \tag{15}$$

using A3) and A5) in the second inequality (observe that the inequality in (13) is reversed in view of A5)). It follows from (15) that R_f is monotone.

ii) For establishing monotonicity of S_f, we can either invoke *i)* and Proposition 2.2, or use (10) with $x_i = y_i$ $(i = 1, 2)$, which holds on account of A2), together with A3) and A4), to obtain

$$-\langle v_1 - v_2, x_1 - x_2 \rangle \leq -f(x_1, x_1) + f(x_1, x_2) + f(x_2, x_1) - f(x_2, x_2) \leq 0,$$

which implies monotonicity of S_f. □

Now we move to the second set of assumptions. We will prove monotonicity of R_f and S_f under A1), A2), A3) and A6). The proof is more involved than that of Theorem 2.3 (we will prove first that R_f and S_f satisfy several properties known to hold for monotone operators, and then get the monotonicity of R_f, S_f as a consequence), and we will motivate it with a special case in which the result is quite immediate, namely the smooth and finite dimensional case.

Proposition 2.4. *Assume that $X = \mathbb{R}^n$ and that f satisfies* A1), A2), A3) *and the following condition stronger than* A6)*: f is continuously differentiable in $X \times X$. Then $R_f = S_f$ and both are monotone.*

Proof. In this case R_f and S_f are point-valued; i.e. $R_f(x) = -\nabla_1 f(x, x)$, $S_f(x) = \nabla_2 f(x, x)$, where ∇_1, ∇_2 have obvious meanings. The Taylor expansion of f gives, for all $(x', y'), (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\begin{aligned} f(x', y') = f(x, y) &+ (\nabla_1 f(x, y), \nabla_2 f(x, y))^t ((x', y') - (x, y)) \\ &+ o\left(\|(x', y') - (x, y)\|\right). \end{aligned} \tag{16}$$

Fix $w \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, and take $y = x$ and $y' = x' = x + \gamma w$, so that (16) becomes

$$f(x', x') = f(x, x) + \gamma (\nabla_1 f(x, x), \nabla_2 f(x, x))^t (w, w)) + o\left(\gamma \|w\|\right). \tag{17}$$

Using A3) in (17),

$$0 = \gamma(\nabla_1 f(x,x) + \nabla_2 f(x,x))^t w + o\left(\gamma \|w\|\right). \tag{18}$$

Dividing (18) by γ and letting $\gamma \to 0$,

$$0 = (\nabla_1 f(x,x) + \nabla_2 f(x,x))^t w = (-R_f(x) + S_f(x))^t w \tag{19}$$

for all $x, w \in \mathbb{R}^n$. It follows from (19) that $R_f(x) = S_f(x)$ for all $x \in \mathbb{R}^n$, and hence

$$R_f = S_f = \frac{1}{2}(R_f + S_f).$$

In view of Proposition 2.1 *ii*), both R_f and S_f are monotone. □

Now we will try to extend the argument in the proof of Proposition 2.4 to the nonsmooth case, but the issue is more delicate, and we will have to establish first several additional properties of R_f, S_f.

We recall that a set-valued operator $T : X \to \mathcal{P}(X^*)$ is *closed-convex-valued* if $T(x)$ is closed and convex for all $x \in X$, and *locally bounded* if for all $x \in X$ there exists a neighborhood U of x such that $\cup_{x\in U} T(x)$ is bounded. Also, the graph $G(T)$ of T is *demiclosed* if for all sequence $\{(x_k, u_k)\} \subset G(T)$ such that $\{x_k\}$ converges strongly to $\bar{x} \in X$ and $\{u_k\}$ converges weakly to $\bar{u} \in X^*$, it holds that $(\bar{x}, \bar{u})$ belongs to $G(T)$.

Proposition 2.5. *If f satisfies* A1) *and* A2) *then*

i) *R_f and S_f are closed-convex-valued.*

ii) *If f also satisfies* A3) *and* A6) *then R_f and S_f are locally bounded.*

iii) *If f also satisfies* A3) *and* A6) *then the graphs of R_f and S_f are demiclosed.*

Proof. We will prove the results only for S_f. Then they will hold also for R_f by virtue of Proposition 2.2.

i) Perform the required elementary computations, or observe that the set $S_f(x)$ is the subdifferential of the convex function $f(x,\cdot)$ evaluated at the point x, and remember that the subdifferential is known to be closed-convex-valued.

ii) Fix $x \in X$. We claim that there exists $\rho > 0$ such that f is bounded in $B(x,\rho) \times B(x,\rho)$. Otherwise there exists a sequence $\{(\hat{z}_k, \tilde{z}_k)\} \subset X \times X$ such that $\lim_{k\to\infty}(\hat{z}_k, \tilde{z}_k) = (0,0)$ and $f(x + \hat{z}_k, x + \tilde{z}_k) \geq 1$ for all k. By A6)

$$f(x,x) = \lim_{k\to\infty} f(x + \hat{z}_k, x + \tilde{z}_k) \geq 1,$$

contradicting A3) and establishing the claim.

Define $\sigma = \rho/2$. We will prove that S_f is bounded on $B(x,\sigma)$. Take θ such that $f(z,y) \leq \theta$ for all $(z,y) \in B(x,\rho) \times B(x,\rho)$. Take any $z \in B(x,\sigma)$ and any nonzero $v \in S_f(z)$.

We will use now the duality operator J. Reflexivity of f implies that J is onto, and that whenever $v \in J(w)$ it holds that $\|v\| = \|w\| = \sqrt{\langle v, w\rangle}$ (see [5]). Use the surjectivity of J to find $w \in J^{-1}(v)$ and take

$$y = z + \frac{\sigma}{\|v\|} w.$$

Note that

$$\|y-x\| \le \|z-x\| + \frac{\sigma}{\|v\|}\|w\| \le \sigma + \sigma = \rho,$$

so that $y \in B(x,\rho)$ and $z \in B(x,\sigma) \subset B(x,\rho)$. Hence $f(z,y) \le \theta$ and therefore, using the fact that $v \in S_f(z)$, the definition of S_f and A3), we get

$$\sigma\|v\| = \frac{\sigma}{\|v\|}\langle v,w\rangle = \langle v, y-z\rangle \le f(z,y) - f(z,z) = f(z,y) \le \theta.$$

It follows that

$$\|v\| \le \frac{\theta}{\sigma}$$

for all $z \in B(x,\sigma)$ and all $v \in S_f(z)$, establishing the local boundedness of S_f.

iii) Take a sequence $\{(x_k,u_k)\} \subset G(S_f)$ such that $\{x_k\}$ is strongly convergent to some $\bar{x} \in X$ and $\{u_k\}$ is weakly convergent to some $\bar{u} \in X^*$. Then, for all $y \in X$,

$$\begin{aligned}\langle \bar{u}, y-\bar{x}\rangle &= \langle \bar{u}-u_k, y-\bar{x}\rangle + \langle u_k, y-\bar{x}\rangle \\ &= \langle \bar{u}-u_k, y-\bar{x}\rangle + \langle u_k, y-x_k\rangle + \langle u_k, x_k-\bar{x}\rangle \\ &\le \langle \bar{u}-u_k, y-\bar{x}\rangle + f(x_k,y) - f(x_k,x_k) + \|u_k\|\,\|\bar{x}-x_k\| \\ &= \langle \bar{u}-u_k, y-\bar{x}\rangle + f(x_k,y) + \|u_k\|\,\|\bar{x}-x_k\|, \end{aligned} \tag{20}$$

using Cauchy-Schwartz inequality, together with the fact that $u_k \in S_f(x_k)$, in the inequality, and A3) in the last equality. Now we take limits with $k \to \infty$ on the rightmost expression of (20). Note that $\lim_{k\to\infty}\langle \bar{u}-u_k, y-\bar{x}\rangle = 0$ by the weak convergence of $\{u_k\}$, $\lim_{k\to\infty}\|\bar{x}-x_k\| = 0$ by the strong convergence of $\{x_k\}$, and $\lim_{k\to\infty} f(x_k,y) = f(\bar{x},y)$ by A6). Also $\{u_k\}$ is bounded as a consequence of *ii*), because the tail of $\{u_k\}$ is contained in $S_f(U)$, for any neighborhood U of $\bar{x}$. It follows that the rightmost expression in (20) converges to $f(\bar{x},y)$ when $k \to \infty$, and hence $\langle \bar{u}, y-\bar{x}\rangle \le f(\bar{x},y) = f(\bar{x},y) - f(\bar{x},\bar{x})$ for all $y \in X$, so that $\bar{u} \in S_f(\bar{x})$ and hence $G(S_f)$ is demiclosed. □

Now we prove monotonicity of both R_f and S_f under our second set of assumptions. The proof of the following theorem can be seen as a nonsmooth version of the proof of Proposition 2.4.

Theorem 2.6. *If f satisfies* A1), A2), A3) *and* A6), *then $S_f = R_f$ and both of them are monotone.*

Proof. We prove first that $S_f(x) \subset R_f(x)$ for all $x \in X$. Fix $x,z \in X$ and define $y_k = x + (1/k)z$. Take $v \in S_f(x), u_k \in R_f(y_k)$. By the definitions of S_f, R_f and A3),

$$\langle v, y_k - x\rangle \le f(y_k,x) - f(x,x) = f(y_k,x), \tag{21}$$

$$\langle -u_k, y_k - x\rangle = \langle u_k, x-y_k\rangle \le -f(y_k,x) - (-f(y_k,y_k)) = -f(y_k,x). \tag{22}$$

Adding (21) and (22),

$$\frac{1}{k}\langle v-u_k, z\rangle = \langle v-u_k, y_k-x\rangle \le 0,$$

implying that

$$\langle v - u_k, z\rangle \leq 0 \tag{23}$$

for all k. Note that $\lim_{k\to\infty} y_k = x$, so that $\{y_k\}$ is bounded, and hence $\{u_k\}$ is bounded by Proposition 2.5 *ii*), because $u_k \in S_f(y_k)$ for all k. Since X is reflexive, it follows from Bourbaki-Alaoglu Theorem (see, e.g. [10], Vol. I, p. 248), that $\{u_k\}$ has weak cluster points; let u be one of them. By Proposition 2.5 *iii*), $u \in S_f(x)$. Taking limits with $k \to \infty$ in (23) along the subsequence which is weakly convergent to u, we get $\langle u - v, z\rangle \geq 0$. We have shown that for all $z \in X$ there exists $u \in R_f(x)$ such that $\langle u - v, z\rangle \geq 0$. Let $V = R_f(x) - v$. By Proposition 2.5 *i*), V is closed and convex, and we have just established that for all $z \in X$ there exists $w \in V$ such that

$$\langle w, z\rangle \geq 0. \tag{24}$$

Invoking again the reflexivity of X, it follows easily from the separation version of Hahn-Banach Theorem (see e.g. Theorem 1.7 in [3]), that $0 \in V$ (otherwise there exists a hiperplane which strictly separates V from 0, contradicting (24)). Now, since $V = R_f(x) - v$, 0 belongs to V if and only if $v \in R_f(x)$. Since v is an arbitrary element of $S_f(x)$, we have proved that $S_f(x) \subset R_f(x)$. The converse inclusion results from Proposition 2.2. It follows that $R_f = S_f = (1/2)(R_f + S_f)$. Monotonicity of R_f and S_f is then a consequence of Proposition 2.1 *ii*). □

3. Maximal monotonicity of the diagonal subdifferential operators

In this section we will prove maximal monotonicity of R_f, S_f under the same assumptions used in Section 2 for establishing their monotonicity, assuming that the space X is such that both the duality operator J and its inverse J^{-1} are single-valued. We remark that single-valuedness of J is equivalent to continuous differentiability of $\|\cdot\|^2$, i.e. to *smoothness* of X. Among Banach spaces satisfying this assumption, we mention the spaces ℓ_p, $\mathcal{L}^p(\Omega)$, and the Sobolev spaces $W^{p,q}(\Omega)$, taking always $1 < p < \infty$.

We need first some preliminary material. We begin with an already mentioned result by Rockafellar.

Proposition 3.1. *Assume that X is a reflexive Banach space such that both J and J^{-1} are single-valued. Let $T : X \to \mathcal{P}(X^*)$ be a monotone operator. If $T + J$ is onto then T is maximal monotone.*

Proof. See Theorem 4.4.7 in [4]. □

It has been proved by E. Asplund that any reflexive Banach space can be equivalently renormed so that J and J^{-1} become single-valued, see [1]. Since this renormalization does not affect the monotonicity, Proposition 3.1 can be used in any reflexive Banach space, with the operator J induced by the new norm.

We continue with a celebrated lemma due to Ky Fan.

Proposition 3.2. *Let Y be a nonempty subset of a real Hausdorff topological vector space Z. Consider a closed-valued $F : Y \to \mathcal{P}(Z)$. If* *i)* *the convex hull of any finite subset* $\{y_1, \ldots, y_m\}$ *of* Y *is contained in* $\bigcup_{i=1}^m F(y_i)$,

ii) *there exists* $y \in Y$ *such that* $F(y)$ *is compact,*

then $\bigcap_{y \in Y} F(y) \neq \emptyset$.

Proof. See Lemma 1 in [6]. □

We remind now that given $f : X \times X \to \mathbb{R}$ and a closed and convex subset $C \subset X$, the *equilibrium problem* EP(f, C) consists of finding $x^* \in C$ such that $f(x^*, x) \geq 0$ for all $x \in C$.

The following property of EP(f, C) appears in [8], with a slightly different formulation.

Proposition 3.3. *Let* K, C *be closed and convex subsets of* X. *Consider a convex* $h : X \to \mathbb{R}$ *and* $f : X \times X \to \mathbb{R}$ *satisfying* A2) *and* A3).

i) *If* $\bar{x}$ *minimizes* h *on* $C \cap K$ *and it belongs to the interior of* K, *then* $\bar{x}$ *minimizes* h *on* C.

ii) *If* $\bar{x}$ *solves EP*$(f, C \cap K)$ *and it belongs to the interior of* K, *then* $\bar{x}$ *solves EP*(f, C).

Proof. Item *i)* is an elementary fact in convex analysis: local minimizers of convex functions are indeed global. We move over to *ii)*. By A2), the marginal function $f_x : X \to \mathbb{R}$ defined as $f_{\bar{x}}(y) = f(\bar{x}, y)$ is convex. Since $\bar{x}$ solves EP$(f, C \cap K)$ we have, in view of A3),

$$f_{\bar{x}}(x) = f(\bar{x}, x) \geq 0 = f(\bar{x}, \bar{x}) = f_{\bar{x}}(\bar{x})$$

for all $x \in C \cap K$, i.e. $\bar{x}$ minimizes the convex function $f_{\bar{x}}$ on $C \cap K$. By *i)*, $\bar{x}$ minimizes $f_{\bar{x}}$ on the whole C, and hence, using again A3),

$$0 = f(\bar{x}, \bar{x}) = f_{\bar{x}}(\bar{x}) \leq f_{\bar{x}}(x) = f(\bar{x}, x)$$

for all $x \in C$, so that $\bar{x}$ solves indeed EP(f, C). □

We will give now a condition on f and C which ensures existence of solutions of EP(f, C) when f satisfies any one of the two set of assumptions considered in Section 2. Several variants of this condition were originally introduced in [8] and further analyzed in [7].

P) For any sequence $\{x_k\} \subset C$ such that $\lim_{k \to \infty} \|x_k\| = \infty$, there exists $w \in C$ and $k_0 \in \mathbb{N}$ such that $f(x_k, w) \leq 0$ for $k \geq k_0$.

We show next that P) guarantees indeed existence of solutions of EP(f, C) under two different sets of assumptions on f.

Theorem 3.4.

i) *If* f *satisfies* A1), A2), A3) *and* P), *then EP*(f, C) *has solutions.*

ii) *If* f *satisfies* A2), A3), A4) *and* P), *and additionally* $f(\cdot, y)$ *is continuous for all* $y \in C$, *then EP*(f, C) *has solutions.*

Proof. *i*) Let C_n be the intersection of C with the ball $B(0,n)$ with radius n centered at 0. Define $F_n : C_n \to \mathcal{P}(X)$ as $F_n(y) = \{x \in C_n : f(x,y) \geq 0\}$. We intend to use Proposition 3.2 for proving existence of solutions of EP(f, C_n), and hence we must check its assumptions. First, we take as Z the Banach space X endowed with its weak topology, under which X is clearly a Haussdorf topological vector space, and $Y = C_n$. The set C_n is certainly closed and in the strong topology of X, and also convex, in view of A1). Hence, it is also weakly closed. We check now assumption *ii*) of Proposition 3.2. In view of A1), $F_n(y)$ is the intersection of C_n with a super-level set of the concave function $f(\cdot, y)$, so that F_n is closed-valued with respect to the topological space Z. We claim now that $F_n(y)$ is compact for all $y \in C$. Note that $F_n(y)$ is weakly closed and also bounded, because it is contained in the bounded set C_n, hence it is weakly compact, i.e. compact in the given topology of Z. Now we must check assumption *i*) of Proposition 3.2. Take $y_1, \dots y_m \in C_n$, and $\alpha_1, \dots, \alpha_m \in \mathbb{R}_+$ such that $\sum_{i=1}^m \alpha_i = 1$. We must verify that $\sum_{i=1}^m \alpha_i y_i \in \bigcup_{i=1}^m F_n(y_i)$, i.e. that there exists ℓ such that

$$f\left(\sum_{i=1}^n \alpha_i y_i, y_\ell\right) \geq 0. \tag{25}$$

Observe that

$$0 = f\left(\sum_{i=1}^m \alpha_i y_i, \sum_{j=1}^m \alpha_j y_j\right) \leq \sum_{j=1}^m \alpha_j f\left(\sum_{i=1}^m \alpha_i y_i, y_j\right) \leq \max_{1\leq j\leq m} f\left(\sum_{i=1}^m \alpha_i y_i, y_j\right), \tag{26}$$

using A3) in the first equality and A2) in the first inequality. Thus, (25) holds if we take as ℓ the index which realizes the maximum in the rightmost expression of (26). The assumptions of Proposition 3.2 therefore hold, and so there exists $x^n \in \bigcap_{y\in C_n} F_n(y)$. It follows from the definition of F_n that $f(x^n, y) \geq 0$ for all $y \in C_n$, and hence x^n solves EP(f, C_n) as claimed. We consider now a sequence $\{x^n\}$ of solutions of EP(f, C_n), whose existence has just been established. We analyze two cases: if there exists n such that $\|x^n\| < n$, then $x^n \in \text{int}(B(0,n))$, and by definition x^n solves EP(f, C_n) =EP$(f, C \cap B(0,n))$. We are thus within the hypotheses of Proposition 3.3 *ii*), and we conclude that x^n solves EP(f, C), establishing the result. We move over to the remaining case, i.e. we assume that $\|x^n\| = n$ for all n, and hence $\lim_{n\to\infty} \|x^n\| = \infty$. Now we invoke assumption P), which ensures the existence of $w \in C$ such that

$$f(x^n, w) \leq 0 \tag{27}$$

for n larger than a given n_0. Take $n > \max\{n_0, \|w\|\}$. Then $w \in C \cap B(0,n) = C_n$, and, since x^n solves EP(f, C_n), it follows that $f(x^n, w) \geq 0$. In view of (27),

$$f(x^n, w) = 0. \tag{28}$$

By (28) and the definition of x^n,

$$f(x^n, w) = 0 \leq f(x^n, y) \quad \forall y \in C_n. \tag{29}$$

Consider the convex function $f_n : X \to \mathbb{R}$, defined as $f_n(y) = f(x^n, y)$. Since w belongs to $C \cap B(0,n)$, w minimizes f_n on $C_n = C \cap B(0,n)$ by (29). Since $\|w\| < n$, w

belongs to the interior of $B(0,n)$. It follows from Proposition 3.3 *i*) that w minimizes f_n on C. Thus, using again (28),

$$0 = f(x^n, w) = f_n(w) \le f_n(y) = f(x^n, y) \quad \forall y \in C,$$

implying that x^n solves EP(f,C), and establishing the result for this case too.

ii) This item has been proved in Theorem 4.2 of [7] under slightly weaker assumptions than those used here, with a proof very similar to the proof of *i*); the main difference lies in the use of monotonicity of f instead of concavity of $f(\cdot, y)$ for establishing that one of the assumptions of Proposition 3.2 holds, namely the weak compacity of $F_n(y)$, which is defined in a different way: $F_n(y) = \{x \in C_n : f(y,x) \le 0\}$. □

The next proposition states that under the remaining assumptions of either item *i*) or item *ii*) of Theorem 3.4, property P) is not only sufficient but also necessary for the existence of solutions of EP(f,C). We need not this result in the sequel, but we deem it interesting enough as to deserve inclusion.

Proposition 3.5. *Assume that f satisfies the assumptions of either item of Theorem* 3.4*, excluding* P)*. If EP(f,C) has solutions then* P) *holds.*

Proof. Let x^* be a solution of EP(f,C). We will show that P) holds with $w = x^*$, and indeed the inequality in P) will hold with any $x \in C$ as the first argument of f, and not just the tail of an unbounded sequence. We consider separately the assumptions of each item of Theorem 3.4. Consider first the hypotheses of item *i*). We have already shown that, due to A2) and A3), the point x^* minimizes the convex function f_{x^*} on C, where $f_{x^*} : X \to \mathbb{R}$ is defined as $f_{x^*}(y) = f(x^*, y)$. By convexity of of f_{x^*} and C, there exists $u^* \in \partial f_{x^*}(x^*) = S_f(x^*)$ such that

$$0 \le \langle u^*, x - x^* \rangle \tag{30}$$

for all $x \in C$. By Theorem 2.6, $S_f = R_f$, so that u^* belongs to R_f. Using (30), the definition of R_f and A3), we get

$$0 \le \langle u^*, x - x^* \rangle \le -f(x, x^*) - (-f(x^*, x^*)) = -f(x, x^*)$$

for all $x \in C$, showing that P) holds indeed with $w = x^*$. The case of item *ii*) was already dealt with in [7], but under A4) the proof is rather immediate:

$$0 \le f(x^*, x) \le -f(x, x^*)$$

for all $x \in C$, using the fact that x^* solves EP(f,C) in the first inequality and A4) in the second one. Again, the inequality in P) holds with $w = x^*$. □

In the theory of equilibrium problems, it is customary to require that assumptions like A1), A2), A3), A4), etc, hold just for points $x, y \in C$ and not in the whole space X. In this case some technical complications arise related to the domains of R_f and S_f. We have opted for a presentation with “unconstrained assumptions” on f just for the sake of clarity of the exposition.

We present now our main result on maximal monotonicity of R_f, S_f.

Theorem 3.6. *Assume that X is a reflexive Banach space. Then,*

i) *if f satisfies* A1), A2), A3) *and* A6), *then $S_f = R_f$ and both of them are maximal monotone,*

ii) *if f satisfies* A2), A3) *and* A4), *and additionally $f(\cdot, y)$ is continuous for all $y \in X$, then S_f is maximal monotone,*

iii) *if f satisfies* A1), A3) *and* A5), *and additionally $f(x, \cdot)$ is continuous for all $x \in X$, then R_f is maximal monotone.*

Proof. In view of Theorems 2.3 and 2.6, it suffices to prove the maximality of S_f, R_f. In view of Proposition 2.2, it suffices to consider the case of S_f, so that we will deal only with items *i*) and *ii*). By Proposition 3.1, it suffices to prove that $S_f + J$ is onto, i.e. that for all $b \in X^*$ there exists $x \in X$ such that $b \in S_f(x) + J(x)$. Define $\tilde{f} : X \times X \to \mathbb{R}$ as

$$\tilde{f}(x, y) = f(x, y) + \frac{1}{2}\left(\|y\|^2 - \|x\|^2\right) + \langle b, x - y\rangle. \tag{31}$$

We assume that J is associated to a norm such that both J and J^{-1} are single-valued, in view of the above mentioned result by E. Asplund.

We will check now that $\tilde{f}$ inherits from f all the properties which appear in the assumptions of either item *i*) or item *ii*), and that EP$(\tilde{f}, X)$ also satisfies P).

Note that $\tilde{f}(x, x) = f(x, x)$ and $\tilde{f}(x, y) + \tilde{f}(y, x) = f(x, y) + f(y, x)$, so that $\tilde{f}$ satisfies A3), A4) or A5) whenever f does. Note also that the second term in the right hand side of (31), namely

$$\frac{1}{2}\left(\|y\|^2 - \|x\|^2\right) + \langle b, x - y\rangle,$$

is convex as a function of y for all $x \in X$, concave as a function of x for all $y \in X$, and jointly continuous as a function of x and y, so that $\tilde{f}$ inherits indeed from f properties A1), A2), A6) and continuity in either argument, when f itself enjoys any of them.

We look now at P) applied to the problem EP$(\tilde{f}, X)$. Let $\{x_k\} \subset X$ be a sequence such that $\lim_{k\to\infty} \|x_k\| = \infty$. Fix any $w \in X$. We will find appropriate upper bounds for $\tilde{f}(x_k, w)$, for which we will consider separately the assumptions of items *i*) and *ii*). We start with item *i*). By A1), $-f(\cdot, w)$ is convex. Let v be a subgradient of this function at w, i.e. v belongs to $R_f(w)$, so that

$$\langle v, x_k - w\rangle \le -f(x_k, w) + f(w, w) = -f(x_k, w),$$

using A3) in the equality, and therefore

$$f(x_k, w) \le \langle v, w - x_k\rangle \le \|v\| \left(\|w\| + \|x_k\|\right). \tag{32}$$

In view of (32) and (31),

$$\begin{aligned}\tilde{f}(x_k,w) &= f(x_k,w)+\frac{1}{2}\left(\|w\|^2-\|x_k\|^2\right)+\langle b,x_k-w\rangle\\ &\le \|v\|\left(\|w\|+\|x_k\|\right)+\frac{1}{2}\left(\|w\|^2-\|x_k\|^2\right)+\|b\|\left(\|x_k\|+\|w\|\right)\\ &= -\frac{1}{2}\|x_k\|^2+\left(\|v\|+\|b\|\right)\|x_k\|+\|w\|\left(\|v\|+\frac{1}{2}\|w\|+\|b\|\right). \qquad (33)\end{aligned}$$

Consider now the assumptions of item *ii*). By A4), $f(x_k,w)\le -f(w,x_k)$. By A3), $f(w,\cdot)$ is convex. Let now v' be a subgradient of this function at w, i.e. $v'\in S_f(w)$, so that

$$\langle v',x_k-w\rangle\le f(w,x_k)-f(w,w)=f(w,x_k),$$

using A3) in the equality, and therefore

$$f(x_k,w)\le -f(w,x_k)\le\langle v',w-x_k\rangle\le\|v'\|\left(\|w\|+\|x_k\|\right). \qquad (34)$$

Proceeding in a similar way from (34) and (31),

$$\tilde{f}(x_k,w)\le -\frac{1}{2}\|x_k\|^2+\left(\|v'\|+\|b\|\right)\|x_k\|+\|w\|\left(\|v'\|+\frac{1}{2}\|w\|+\|b\|\right), \qquad (35)$$

Since $\lim_{k\to\infty}\|x_k\|=\infty$, it follows from either (33) or (35) that $\lim_{k\to\infty}\tilde{f}(x_k,w)=-\infty$, so that $\tilde{f}(x_k,w)\le 0$ for large enough k, and hence P) holds under the assumptions of either item *i*) or item *ii*).

We have checked all the required properties of $\tilde{f}$, so that we can apply Theorem 3.4, in order to conclude that $\mathrm{EP}(\tilde{f},X)$ has solutions, under the hypotheses of either item *i*) or item *ii*). Let x^* be a solution of $\mathrm{EP}(\tilde{f},X)$, i.e., since $\tilde{f}$ satisfies A3), it holds that $\tilde{f}(x^*,x^*)=0\le\tilde{f}(x^*,x)$ for all $x\in X$, so that x^* is an unrestricted minimizer of the convex function $\tilde{f}(x^*,\cdot)$, and hence a zero of its subdifferential at x^*, namely $S_{\tilde{f}}(x^*)$. It follows easily from(31) and the definition of J that $S_{\tilde{f}}(x)=S_f(x)+J(x)-b$ for all $x\in X$. We have proved that $0\in(S_f+J)(x^*)-b$, i.e. that $b\in(S_f+J)(x^*)$. Since b is an arbitrary element of X^*, it follows that S_f+J is onto, and hence S_f is maximal monotone by Proposition 3.1. The proof is complete. □

We mention that the result of Theorem 3.6 *ii*) has been proved, for the special case in which X is a Hilbert space, in [9], but with another another regularization function $\hat{f}$, defined as $\hat{f}(x,y)=f(x,y)+\langle\lambda x-b,y-x\rangle$, instead of $\tilde{f}$. We remark that this function $\hat{f}$ cannot be adequately extended to Banach spaces.

We conjecture that S_f and R_f are maximal monotone in any Banach space X under the remaining assumptions of Theorem 3.6.

References

[1] E. Asplund: Averaged norms, Israel J. Math. 5 (1967) 227–233.

[2] E. Blum, W. Oettli: From optimization and variational inequalities to equilibrium problems, Math. Stud. 63 (1994) 123–145.

[3] H. Brezis: Analyse Fonctionnelle. Théorie et Applications, Masson, Paris (1983).

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

[5] J. Diestel: Geometry of Banach Spaces. Selected Topics, Springer, Berlin (1975).

[6] K. Fan: A generalization of Tychonoff's fixed point theorem, Math. Ann. 142 (1961) 305–310.

[7] A. N. Iusem, G. Kassay, W. Sosa: On certain conditions for the existence of solutions of equilibrium problems, Math. Program., Ser. B 116 (2009) 259–273.

[8] A. N. Iusem, W. Sosa: New existence results for equilibrium problems, Nonlinear Anal. 52 (2003) 621–635.

[9] A. N. Iusem, W. Sosa: A proximal point method for equilibrium problems in Hilbert spaces, Optimization, to appear.

[10] G. Köthe: Topological Vector Spaces, Springer, New York (1969).