Journal of Nonlinear and Convex Analysis
Volume 6, Number 1, 2005, 1–22
Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

FENCHEL DUALITY, FITZPATRICK FUNCTIONS AND MAXIMAL MONOTONICITY

S. SIMONS AND C. ZĂLINESCU

This paper is dedicated to the memory of Simon Fitzpatrick, in recognition of his amazing insights

Abstract. We show in this paper how the versions of the Fenchel duality theorem due to Rockafellar and Attouch–Brezis can be applied to the Fitzpatrick function determined by a maximal monotone multifunction to obtain a number of results on maximal monotonicity, including a number of sufficient conditions for the sum of maximal monotone multifunctions on a reflexive Banach space to be maximal monotone, unifying a number of the results of "Attouch–Brezis type" that have been obtained in recent years. We also obtain generalizations of the Brezis–Crandall–Pazy result. We find various explicit formulas in terms of the Fitzpatrick function for the minimum norm of the solutions x of $(S+J)x \ni 0$, where E is reflexive, S is maximal monotone on E and J is the duality map. Among the tools that we develop are a version of the Fenchel duality theorem in which we obtain an explicit formula for the minimum norm of solutions in certain cases, and a generalization of the Attouch–Brezis version of the Fenchel duality theorem to a more symmetric result for convex functions of two variables.

0. Introduction

We start off by stating a result that is an immediate consequence of Rockafellar's version of the Fenchel duality theorem (see [12, Theorem 1, p. 82–83] for the original version and [16, Theorem 2.8.7, p. 126–127] for more general results):

Theorem 0.1. *Let F be a normed space, $f\colon\ F \mapsto (-\infty,\infty]$ be proper and convex, $g\colon\ F \mapsto \mathbb{R}$ be convex and continuous, and $f+g \geq \lambda$ on F. Then there exists $x^* \in F^*$ such that $f^*(x^*)+g^*(-x^*) \leq -\lambda$.*

We show in this paper how Theorem 0.1 and the Attouch–Brezis version of the Fenchel duality theorem (see Theorem 4.1 below) can be used to obtain a number of results on maximal monotonicity, including a number of sufficient conditions for the sum of maximal monotone multifunctions on a reflexive Banach space to be maximal monotone.

In Section 1, we show how certain convex functions on $E \times E^*$ (E reflexive) lead to graphs of maximal monotone multifunctions. The main results here are Lemma 1.2(c), which will be used in our work on the Brezis–Crandall–Pazy condition for the sum of maximal monotone multifunctions to be maximal monotone (see Theorem 6.2), and Theorem 1.4, which generalizes a result proved by Burachik and Svaiter in [5] (see the discussion preceding Theorem 1.4 for more details of this).

2000 *Mathematics Subject Classification.* 47H05, 46B10, 46N10.

Key words and phrases. Monotone multifunction, Banach space, Convex analysis, Fenchel duality theorem.

Section 2 is devoted to a single result, Theorem 2.1. Here we bootstrap Theorem 0.1 in the special case where $g(x) := \frac{1}{2}\|x\|^2$ and $\lambda := 0$ to find a sharp lower bound on the norm of the functionals x^* that satisfy the conclusion of Theorem 0.1. This lower bound will be used in Theorem 3.1 to find the minimum norm of the solutions x of $(S + J)x \ni 0$, where E is reflexive, S is maximal monotone on E and J is the duality map.

Our results on the maximal monotonicity of a sum use the Fitzpatrick function determined by a maximal monotone multifunction. The elementary properties of this function will be explained in Section 3. The main result in this section is Theorem 3.1, which we have already discussed above, and which will be used in our work on the Brezis–Crandall–Pazy condition (see Lemma 6.1).

In Theorem 4.2, we show how the Attouch–Brezis version of the Fenchel duality theorem can be generalized to a more symmetric version for convex functions of two variables.

We give in Theorem 5.5 a sufficient condition for the sum of maximal monotone multifunctions on a reflexive Banach space to be maximal monotone, unifying a number of the results of "Attouch–Brezis type" that have been obtained in recent years. In order to do this, we start off by combining the results of Theorem 4.2 and Theorem 1.4(a) to establish a special case in Lemma 5.1, and then bootstrapping Lemma 5.1 with a sequence of three lemmas in order to obtain Theorem 5.5. We mention parenthetically that we use Theorem 1.4(a) rather than Theorem 1.4(b) in Lemma 5.1 since we do not know that the function ρ is lower semicontinuous.

In Section 6, we use Theorem 3.1(a) and Lemma 1.2(c) to obtain generalizations of the Brezis–Crandall–Pazy result.

In the final section, we give alternative formulas for the minimum norm of the solutions x of $(S + J)x \ni 0$ already discussed in Theorem 3.1.

We close this introduction with some remarks of a more historical nature. The first person to apply convex analysis to the representation of maximal monotone multifunctions was Krauss ([7]), who gave a represention in terms of saddle functions. Fitzpatrick ([6]) was the first person to give a representation in terms of convex functions. Fitzpatrick's results were rediscovered by Burachik–Svaiter ([4]) and Martínez-Legaz–Théra ([8]). The first people to use Fitzpatrick functions to obtain a proof of Rockafellar's surjectivity theorems or sufficient conditions for the sum of maximal monotone multifunctions (for reflexive spaces) were Zălinescu and Penot. Indeed, the authors would like to thank Jean-Paul Penot for sending them copies of [9], which was a considerable source of inspiration.

1. Convex functions on $E \times E^*$ for reflexive E

In this section, we assume that E is a reflexive Banach space and E^* is its topological dual space. We norm $E \times E^*$ by $\|(x, x^*)\| := \sqrt{\|x\|^2 + \|x^*\|^2}$. Then the topological dual of $E \times E^*$ is $E^* \times E$, under the pairing $\langle (x, x^*), (u^*, u)\rangle := \langle x, u^*\rangle + \langle u, x^*\rangle$. Further, $\|(u^*, u)\| = \sqrt{\|u\|^2 + \|u^*\|^2}$.

Notation 1.1. In order to simplify some rather cumbersome algebraic expressions, we will define $\Delta\colon\ E \times E^* \mapsto \mathbb{R}$ by $\Delta(y, y^*) := \langle y, y^*\rangle + \frac{1}{2}\|(y, y^*)\|^2$. "$\Delta$" stands for "discriminant". We note then that, for all $(y, y^*) \in E \times E^*$,

$$\Delta(y, y^*) = \tfrac{1}{2}\|y\|^2 + \langle y, y^*\rangle + \tfrac{1}{2}\|y^*\|^2 \geq \tfrac{1}{2}\|y\|^2 - \|y\|\|y^*\| + \tfrac{1}{2}\|y^*\|^2 \geq 0. \tag{1.1.1}$$

Clearly $\Delta(y, y^*) = 0 \Longrightarrow \|y^*\| = \|y\|$. Plugging this back into (1.1.1), we have

$$\Delta(y, y^*) = 0 \Longrightarrow \langle y, y^*\rangle = -\|y\|^2 = -\|y^*\|^2 = -\|y\|\|y^*\|. \tag{1.1.2}$$

The significance of this is that, if $J\colon E \rightrightarrows E^*$ is the duality map, then

$$\Delta(y, y^*) = 0 \iff y^* \in -Jy. \tag{1.1.3}$$

Lemma 1.2. *Let $h\colon\ E \times E^* \mapsto (-\infty, \infty]$ be convex and*

$$(x, x^*) \in E \times E^* \Longrightarrow h(x, x^*) \geq \langle x, x^*\rangle. \tag{1.2.1}$$

Write M_h for the set $\big\{(x, x^*) \in E \times E^*\colon\ h(x, x^*) = \langle x, x^*\rangle\big\}$.

(a) *M_h is a monotone subset of $E \times E^*$.*

(b) *Let (w, w^*) and $(x, x^*) \in E \times E^*$ be such that*

$$h(x, x^*) - \langle x, x^*\rangle + \Delta(w - x, w^* - x^*) \leq 0. \tag{1.2.2}$$

Then $(x, x^*) \in M_h$.

(c) *Suppose that $G \subset M_h$ and, for all $(w, w^*) \in E \times E^*$ there exists $(x, x^*) \in G$ satisfying (1.2.2). Then G is a maximal monotone subset of $E \times E^*$ (and consequently, $G = M_h$).*

Proof. (a) Let $(x, x^*), (y, y^*) \in M_h$. Then, from the convexity of h and (1.2.1),

$$\begin{aligned}\tfrac{1}{2}\langle x, x^*\rangle + \tfrac{1}{2}\langle y, y^*\rangle &= \tfrac{1}{2}h(x, x^*) + \tfrac{1}{2}h(y, y^*)\\ &\geq h(\tfrac{1}{2}x + \tfrac{1}{2}y, \tfrac{1}{2}x^* + \tfrac{1}{2}y^*) \geq \langle \tfrac{1}{2}x + \tfrac{1}{2}y, \tfrac{1}{2}x^* + \tfrac{1}{2}y^*\rangle.\end{aligned}$$

This implies that $\langle x - y, x^* - y^*\rangle \geq 0$, and so M_h is monotone.

(b) (1.2.2) and (1.1.1) give $h\big(x, x^*\big) \leq \langle x, x^*\rangle$, and it is clear from (1.2.1) that $(x, x^*) \in M_h$.

(c) Since $G \subset M_h$, it follows from (a) that G is monotone. In order to prove that G is maximal monotone, we suppose that $(w, w^*) \in E \times E^*$ and

$$(x, x^*) \in G \quad\Longrightarrow\quad \langle w - x, w^* - x^*\rangle \geq 0, \tag{1.2.3}$$

(i.e., (w, w^*) is "monotonically related" to G) and we will deduce that

$$(w, w^*) \in G. \tag{1.2.4}$$

To this end, we choose $(x, x^*) \in G$ as in (1.2.2). Using (1.2.1), we derive from this that $\Delta(w - x, w^* - x^*) \leq 0$ thus, from (1.2.3), $\tfrac{1}{2}\big\|(w - x, w^* - x^*)\big\|^2 \leq 0$, so $(w, w^*) = (x, x^*) \in G$. This establishes (1.2.4), and completes the proof of (c). □

Lemma 1.3. *Let E be a reflexive Banach space, $k\colon\ E \times E^* \mapsto (-\infty, \infty]$ be proper and convex, $(w, w^*) \in E \times E^*$ and*

$$(x, x^*) \in E \times E^* \quad\Longrightarrow\quad k(x, x^*) - \langle x, x^*\rangle + \Delta(w - x, w^* - x^*) \geq 0. \tag{1.3.1}$$

Then

$$\text{there exists } (x, x^*) \in E \times E^* \text{ such that } k^*(x^*, x) - \langle x, x^*\rangle + \Delta(w - x, w^* - x^*) \leq 0. \tag{1.3.2}$$

Proof. Define $\delta_{(w,w^*)}\colon\ E \times E^* \mapsto \mathbb{R}$ by

$$\delta_{(w,w^*)}(x,x^*) := -\langle x,x^*\rangle + \Delta(w-x,w^*-x^*). \tag{1.3.3}$$

Then the identity $\delta_{(w,w^*)}(x,x^*) = \langle w,w^*\rangle - \big\langle (x,x^*),(w^*,w)\big\rangle + \frac{1}{2}\big\|(w,w^*)-(x,x^*)\big\|^2$ shows that

(1.3.4) $\delta_{(w,w^*)}$ is convex and norm–continuous,

hence weakly lower semicontinuous.

(The weak lower semicontinuity will be used in Theorem 6.2.) By direct computation,

$$(x,x^*) \in E \times E^* \quad\Longrightarrow\quad {\delta_{(w,w^*)}}^*(-x^*,-x) = \delta_{(w,w^*)}(x,x^*). \tag{1.3.5}$$

Now (1.3.1) gives $\inf_{E\times E^*}\big[k + \delta_{(w,w^*)}\big] \ge 0$; thus we can deduce from Theorem 0.1 that $\min_{\eta^*\in E^*\times E}\big[k^*(\eta^*) + {\delta_{(w,w^*)}}^*(-\eta^*)\big] \le 0$. Consequently, (1.3.2) now follows from (1.3.5). □

Theorem 1.4(b) below was first established in [5, Theorem 3.1]. The proof given here avoids having to use a renorming theorem. The interest of Theorem 1.4(a) is that the function k is not required to be lower semicontinuous, which fact will be very useful to us in Lemma 5.1. In fact, Theorem 1.4(a) can be deduced from Theorem 1.4(b) using a technique similar to that of [9, Theorem 15].

Theorem 1.4. (a) *Let $k\colon\ E \times E^* \mapsto (-\infty,\infty]$ be proper and convex,*

$$(x,x^*) \in E \times E^* \Longrightarrow k(x,x^*) \ge \langle x,x^*\rangle. \tag{1.4.1}$$

and

$$(x,x^*) \in E \times E^* \Longrightarrow k^*(x^*,x) \ge \langle x,x^*\rangle. \tag{1.4.2}$$

Then $G := \big\{(x,x^*) \in E \times E^*\colon\ k^*(x^*,x) = \langle x,x^*\rangle\big\}$ is a maximal monotone subset of $E \times E^*$.
(b) *Let $h\colon\ E \times E^* \mapsto (-\infty,\infty]$ be proper, convex and lower semicontinuous,*

$$(x,x^*) \in E \times E^* \Longrightarrow h(x,x^*) \ge \langle x,x^*\rangle. \tag{1.2.1}$$

and

$$(x,x^*) \in E \times E^* \Longrightarrow h^*(x^*,x) \ge \langle x,x^*\rangle. \tag{1.4.3}$$

Then $M_h := \big\{(x,x^*) \in E \times E^*\colon\ h(x,x^*) = \langle x,x^*\rangle\big\}$ is a maximal monotone subset of $E \times E^*$.

Proof. (a) Let (w,w^*) be an arbitrary element of $E \times E^*$. Then (1.3.1) follows from (1.4.1) and (1.1.1), and so Lemma 1.3 gives (1.3.2). Combining this with (1.1.1) and (1.4.2), we have $k^*(x^*,x) = \langle x,x^*\rangle$, that is to say $(x,x^*) \in G$, and (1.2.2) is satisfied with $h(x,x^*) := k^*(x^*,x)$. (a) now follows from Lemma 1.2.

(b) Let $k(x,x^*) := h^*(x^*,x)$. k is proper and convex on $E \times E^*$ and (1.4.1) follows from (1.4.3). Since h is lower semicontinuous, the Fenchel–Moreau formula shows that $k^*(x^*,x) = h(x,x^*)$, and so (1.4.2) follows from (1.2.1). (b) is now immediate from (a). □

2. Fenchel duality with a sharp lower bound on the norm

It is immediate from Theorem 0.1 that (2.1.1) below implies the existence of $x^* \in F^*$ satisfying (2.1.3). Theorem 2.1(a) gives the additional information that there exists such a functional x^* with $\|x^*\| \le M$. This information will be used in Theorem 3.1 and Lemma 6.1. Theorem 2.1(b) shows that this value of M is best possible. Of course, the crux of the proof of Theorem 2.1 is the advance knowledge of the "magic number" M. In what follows, for all $\lambda \in \mathbb{R}$ we write λ^+ for $\lambda \vee 0$.

Theorem 2.1. (a) *Let F be a normed space, $f\colon\ F \mapsto (-\infty, \infty]$ be proper and convex and*

$$x \in F \quad \Longrightarrow \quad f(x) + \tfrac{1}{2}\|x\|^2 \ge 0. \tag{2.1.1}$$

Let

$$M := \sup_{x \in F} \Big[\|x\| - \sqrt{2f(x) + \|x\|^2}\Big]^+. \tag{2.1.2}$$

Then there exists $x^* \in F^*$ such that $\|x^*\| \le M$ and

$$f^*(x^*) + \tfrac{1}{2}\|x^*\|^2 \le 0. \tag{2.1.3}$$

Further,

$$M \le \inf_{x \in F} \Big[\|x\| + \sqrt{2f(x) + \|x\|^2}\Big]. \tag{2.1.4}$$

(b) *If $x^* \in F^*$ satisfies* (2.1.3), *then $\|x^*\| \ge M$.*

Proof. We observe from (2.1.1) that the square root in (2.1.2) is real (or $+\infty$). We start off by showing that

$$u, v \in F \quad \Longrightarrow \quad \|v\| - \sqrt{2f(v) + \|v\|^2} \le \|u\| + \sqrt{2f(u) + \|u\|^2}. \tag{2.1.5}$$

To this end, let $u, v \in F$. We can clearly suppose that $f(u) \in \mathbb{R}$ and $f(v) \in \mathbb{R}$. Let $\lambda > \sqrt{2f(u) + \|u\|^2} \ge 0$ and $\mu > \sqrt{2f(v) + \|v\|^2} \ge 0$, and write $\alpha := \|u\| + \lambda$ and $\beta := \|v\| - \mu$. Then, since $\mu\|u\| + \lambda\|v\| = \mu\alpha + \lambda\beta$,

$$0 \le \Big\|\frac{\mu u + \lambda v}{\mu + \lambda}\Big\| \le \frac{\mu\|u\| + \lambda\|v\|}{\mu + \lambda} = \frac{\mu\alpha + \lambda\beta}{\mu + \lambda}.$$

Thus, from (2.1.1) applied to $x = \dfrac{\mu u + \lambda v}{\mu + \lambda} \in F$, and the convexity of f and $(\cdot)^2$,

$$\begin{aligned}\frac{\mu f(u) + \lambda f(v)}{\mu + \lambda} \ge f\Big(\frac{\mu u + \lambda v}{\mu + \lambda}\Big) &\ge -\frac{1}{2}\Big\|\frac{\mu u + \lambda v}{\mu + \lambda}\Big\|^2 \\ &\ge -\frac{1}{2}\Big(\frac{\mu\alpha + \lambda\beta}{\mu + \lambda}\Big)^2 \ge -\frac{1}{2}\frac{\mu\alpha^2 + \lambda\beta^2}{\mu + \lambda}.\end{aligned}$$

Multiplying by $2(\mu + \lambda)$ gives

$$\begin{aligned}0 &\le 2\mu f(u) + 2\lambda f(v) + \mu\alpha^2 + \lambda\beta^2 = \mu\big(2f(u) + \alpha^2\big) + \lambda\big(2f(v) + \beta^2\big) \\ &= \mu\big(2f(u) + \|u\|^2 + 2\lambda\|u\| + \lambda^2\big) + \lambda\big(2f(v) + \|v\|^2 - 2\mu\|v\| + \mu^2\big) \\ &< \mu\big(2\lambda^2 + 2\lambda\|u\|\big) + \lambda\big(2\mu^2 - 2\mu\|v\|\big) = 2\mu\lambda\big(\lambda + \|u\| + \mu - \|v\|\big).\end{aligned}$$

On dividing by $2\mu\lambda$, we obtain $\|v\| - \mu < \|u\| + \lambda$, and (2.1.5) follows by letting $\mu \to \sqrt{2f(v) + \|v\|^2}$ and $\lambda \to \sqrt{2f(u) + \|u\|^2}$. Now (2.1.2) and (2.1.5) imply that, for all $x \in F$,

$$\|x\| - \sqrt{2f(x) + \|x\|^2} \le M \quad \text{and} \quad M \le \|x\| + \sqrt{2f(x) + \|x\|^2}, \tag{2.1.6}$$

from which

$$\begin{aligned} x \in F \quad &\Longrightarrow \quad \big|\|x\| - M\big| \le \sqrt{2f(x) + \|x\|^2} \\ &\Longrightarrow \quad \big(\|x\| - M\big)^2 \le 2f(x) + \|x\|^2 \\ &\Longrightarrow \quad f(x) + M\|x\| \ge \tfrac{1}{2}M^2. \end{aligned}$$

Theorem 0.1 now gives the existence of $x^* \in F^*$ such that $f^*(x^*) + (M\|\cdot\|)^*(-x^*) \le -\frac{1}{2}M^2$, thus $\|x^*\| \le M$ and $f^*(x^*) \le -\frac{1}{2}M^2$, from which (2.1.3) is immediate. Since (2.1.6) implies (2.1.4), this completes the proof of (a).

(b) Now suppose that $x^* \in F^*$ satisfies (2.1.3), and let x be an arbitrary element of F. It follows from (2.1.3) that

$$f(x) \ge \langle x, x^*\rangle - f^*(x^*) \ge \langle x, x^*\rangle + \tfrac{1}{2}\|x^*\|^2 \ge -\|x\|\|x^*\| + \tfrac{1}{2}\|x^*\|^2,$$

and so $2f(x) + \|x\|^2 \ge \|x\|^2 - 2\|x\|\|x^*\| + \|x^*\|^2 = \big(\|x\| - \|x^*\|\big)^2$. Thus

$$\sqrt{2f(x) + \|x\|^2} \ge \|x\| - \|x^*\|,$$

from which $\|x^*\| \ge \|x\| - \sqrt{2f(x) + \|x\|^2}$, and (b) follows by taking the supremum over $x \in F$. □

3. The Fitzpatrick function and surjectivity

Let E be a reflexive Banach space and $S\colon\ E \rightrightarrows E^*$ be maximal monotone with graph $G(S) := \big\{(x, x^*) \in E \times E^*\colon\ x^* \in Sx\big\}$. We define $\psi_S\colon\ E \times E^* \mapsto (-\infty, \infty]$ by

$$\psi_S(x, x^*) := \sup_{(s,s^*) \in G(S)} \langle x - s, s^* - x^*\rangle,$$

and the *Fitzpatrick function* $\phi_S\colon\ E \times E^* \mapsto (-\infty, \infty]$ associated with S by

$$\phi_S(x, x^*) := \sup_{(s,s^*) \in G(S)} \big[\langle s, x^*\rangle + \langle x, s^*\rangle - \langle s, s^*\rangle\big] = \psi_S(x, x^*) + \langle x, x^*\rangle.$$

(This function ϕ_S was introduced by Fitzpatrick in [6, Definition 3.1, p. 61] under the notation L_S. The function ψ_S was introduced by Brezis and Haraux in [3] and used further by Reich in [10] in their work on the range of the sum of monotone multifunctions on a reflexive Banach space.) The reader may ask why we have introduced both the functions ϕ_S and ψ_S, which are so closely related. The reason for this is that ϕ_S is convex and weakly lower semicontinuous, while ψ_S is generally neither. On the other hand, ψ_S is positive while ϕ_S is generally not. So the choice of which of the two functions we use depends on what kind of argument we are employing. A good example of this can be found in the transition from the use of ψ_S and ψ_T in (6.2.3) to the use of ϕ_S and ϕ_T in (6.2.5). Now the maximal monotonicity of S gives the statements

$$(x, x^*) \in E \times E^* \quad \Longrightarrow \quad \psi_S(x, x^*) \ge 0 \iff \phi_S(x, x^*) \ge \langle x, x^*\rangle, \tag{3.0.1}$$

and

$$\psi_S(x,x^*)=0 \iff \phi_S(x,x^*)=\langle x,x^*\rangle \iff (x,x^*)\in G(S). \tag{3.0.2}$$

(See [6, Corollary 3.9, p. 62]. In fact, the monotonicity of S implies that if $(x,x^*)\in G(S)$ then $\psi_S(x,x^*)=0$, while the maximality of S implies that if $(x,x^*)\in E\times E^*\setminus G(S)$ then $\psi_S(x,x^*)>0$.) We will have frequent occasion to use the identity, immediate from (3.0.1), that

$$\eta\in E\times E^* \implies \phi_S(\eta)+\tfrac{1}{2}\|\eta\|^2=\psi_S(\eta)+\Delta(\eta)\geq 0. \tag{3.0.3}$$

Taken together with (3.0.3) and (1.1.3), (3.0.2) implies that

$$\phi_S(\eta)+\tfrac{1}{2}\|\eta\|^2=0 \iff \psi_S(\eta)+\Delta(\eta)=0 \iff \eta\in G(S)\cap G(-J). \tag{3.0.4}$$

Clearly, ϕ_S is proper, convex and lower semicontinuous. Let $(x,x^*)\in E\times E^*$. Then we see from the definition of ϕ_S and (3.0.2) that

$$\begin{aligned}\phi_S(x,x^*)&=\sup_{(s,s^*)\in G(S)}\big[\langle s,x^*\rangle+\langle x,s^*\rangle-\phi_S(s,s^*)\big]\\&\leq\sup_{(y,y^*)\in E\times E^*}\big[\langle y,x^*\rangle+\langle x,y^*\rangle-\phi_S(y,y^*)\big]\\&=\sup_{(y,y^*)\in E\times E^*}\big[\big\langle (y,y^*),(x^*,x)\big\rangle-\phi_S(y,y^*)\big]=\phi_S^*(x^*,x).\end{aligned}$$

Combining this with (3.0.1), we have (see [6, Proposition 4.2, p. 63])

$$(x,x^*)\in E\times E^* \implies \phi_S^*(x^*,x)\geq\phi_S(x,x^*)\geq\langle x,x^*\rangle. \tag{3.0.5}$$

Further, if $(x,x^*)\in G(S)$ then, for all $(y,y^*)\in E\times E^*$, the definition of $\phi_S(y,y^*)$ yields

$$\phi_S(y,y^*)\geq\langle y,x^*\rangle+\langle x,y^*\rangle-\langle x,x^*\rangle=\big\langle (y,y^*),(x^*,x)\big\rangle-\langle x,x^*\rangle.$$

Thus

$$\begin{aligned}&(x,x^*)\in G(S)\\&\qquad\implies \phi_S^*(x^*,x)=\sup_{(y,y^*)\in E\times E^*}\big[\big\langle (y,y^*),(x^*,x)\big\rangle-\phi_S(y,y^*)\big]\leq\langle x,x^*\rangle.\end{aligned}$$

Combining this with (3.0.2) and (3.0.5) yields (see [6, Proposition 4.3, p. 63])

$$\phi_S^*(x^*,x)=\langle x,x^*\rangle \iff (x,x^*)\in G(S). \tag{3.0.6}$$

We now show how ϕ_S can be used to establish Rockafellar's surjectivity theorem that $R(S+J)\ni 0$ and give a sharp lower bound in terms of ϕ_S for the norm of the solutions, s, of $(S+J)s\ni 0$. This can, of course, be bootstrapped into a proof that $E\times E^*=G(S)+G(-J)$ (see [13, Theorem 10.6, p. 37] and [15, Theorem 1.2]), with the appropriate sharp numerical estimates. The numerical estimates obtained in Theorem 3.1 will be used in Lemma 6.1. We emphasize that we have not assumed that E has been renormed in any particular way.

Theorem 3.1. *Let E be a nontrivial reflexive Banach space, $S\colon E\rightrightarrows E^*$ be a maximal monotone multifunction and*

$$N:=\tfrac{1}{\sqrt{2}}\sup\nolimits_{\eta\in E\times E^*}\Big[\|\eta\|-\sqrt{2\phi_S(\eta)+\|\eta\|^2}\Big]^+.$$

(a) *There exists* $\eta^* \in E^* \times E$ *such that* $\|\eta^*\| \le \sqrt{2}N$ *and*

$$\phi_S^*(\eta^*) + \tfrac{1}{2}\|\eta^*\|^2 \le 0. \tag{3.1.1}$$

Let $(z, z^*) \in E \times E^*$ *be such that* $\eta^* = (z^*, z)$. *Then*

$$\tfrac{1}{2}\big\|(z, z^*)\big\|^2 \le N^2 \tag{3,1,2}$$

and

$$\phi_S(z, z^*) + \tfrac{1}{2}\big\|(z, z^*)\big\|^2 = \psi_S(z, z^*) + \Delta(z, z^*) \le 0. \tag{3.1.3}$$

Finally,

$$\begin{cases} N \le \frac{1}{\sqrt{2}} \inf_{\eta \in E \times E^*} \Big[\|\eta\| + \sqrt{2\phi_S(\eta) + \|\eta\|^2}\Big] \\ \quad = \frac{1}{\sqrt{2}} \inf_{\eta \in E \times E^*} \Big[\|\eta\| + \sqrt{2\psi_S(\eta) + 2\Delta(\eta)}\Big]. \end{cases} \tag{3.1.4}$$

(b) *There exists* $x \in E$ *such that* $(S + J)x \ni 0$, *and further*

$$\min\big\{\|x\|\colon\ x \in E,\ (S + J)x \ni 0\big\} = N.$$

Proof. (a) It is immediate from (3.0.3) and Theorem 2.1(a) with $F := E \times E^*$ and $f := \phi_S$ that there exists $\eta^* \in E^* \times E$ satisfying (3.1.1) such that $\|\eta^*\| \le \sqrt{2}N$. (3,1,2) is also clear since $\big\|(z, z^*)\big\|^2 = \|\eta^*\|^2$. (3.1.3) now follows from (3.1.1), (3.0.5) and (3.0.3), and (3.1.4) follows from (2.1.4). This completes the proof of (a).

(b) If (z, z^*) is as in (a), then (3.1.3), (3.0.3) and (**??**) give us that $(z, z^*) \in G(S)$ and $-z^* \in Jz$. Since $0 = z^* + (-z^*)$, it is now immediate that $(S + J)z \ni 0$, and (1.1.2) and (3,1,2) imply that $\|z\| \le N$. In order to complete the proof of (b), we must show that

$$x \in E \text{ and } (S + J)x \ni 0 \quad \Longrightarrow \quad \|x\| \ge N \tag{3.1.5}$$

So suppose that $x \in E$ and $(S + J)x \ni 0$. Then there exists $x^* \in Sx$ such that $-x^* \in Jx$. From (3.0.6), $\phi_S^*(x^*, x) + \frac{1}{2}\big\|(x^*, x)\big\|^2 = \langle x, x^*\rangle + \frac{1}{2}\big\|(x^*, x)\big\|^2 = \frac{1}{2}\|x\|^2 - \|x\|\|x^*\| + \frac{1}{2}\|x^*\|^2 = \frac{1}{2}\big(\|x\| - \|x^*\|\big)^2 = 0$. Theorem 2.1(b) now gives $\sqrt{2}\|x\| = \big\|(x^*, x)\big\| \ge \sqrt{2}N$, from which (3.1.5) follows, completing the proof of (b). □

Remark 3.2. Some of the techniques introduced in this section have been used in [11] to give a proof of the Kirszbraun–Valentine extension theorem for nonexpansive maps on a Hilbert space.

4. A More symmetric version of a result of Attouch and Brezis

For the initial results of this section we consider (possibly) nonreflexive Banach spaces. Theorem 4.1 below was first proved by Attouch–Brezis (this follows from [1, Corollary 2.3, p. 131–132]) — there is a somewhat different proof in [13, Theorem 14.2, p. 51], and a much more general result was established in [16, Theorem 2.8.6, p. 125–126]:

Theorem 4.1. *Let* K *be a Banach space,* $f,\ g\colon\ K \mapsto (-\infty, \infty]$ *be convex and lower semicontinuous,*

$$\bigcup\nolimits_{\lambda > 0} \lambda\big[\,\mathrm{dom}\ f - \mathrm{dom}\ g\big] \ \textit{be a closed subspace of } K$$

and

$$f+g\geq 0 \ \text{on}\ K.$$

Then

$$\text{there exists } z^*\in K^* \text{ such that}\quad f^*(-z^*)+g^*(z^*)\leq 0.$$

Our next result is a generalization of Theorem 4.1 to functions of two variables. We note that $\rho(x,\cdot)$ is the inf–convolution (=episum) of $\sigma(x,\cdot)$ and $\tau(x,\cdot)$, and the conclusion of Theorem 4.2 is that $\rho^*(\cdot,y^*)$ is the exact inf–convolution of $\sigma^*(\cdot,y^*)$ and $\tau^*(\cdot,y^*)$.

Theorem 4.2. *Let E and F be Banach spaces, $\sigma,\ \tau\colon\ E\times F\mapsto(-\infty,\infty]$ be proper, convex and lower semicontinuous and, for all $(x,y)\in E\times F$,*

$$\rho(x,y):=\inf\big\{\sigma(x,u)+\tau(x,v)\colon\ u,\,v\in F,\ u+v=y\big\}>-\infty.$$

Defining $\mathrm{pr}_1(x,y):=x$, let

$$L:=\bigcup\nolimits_{\lambda>0}\lambda\big[\,\mathrm{pr}_1\ \mathrm{dom}\ \sigma-\mathrm{pr}_1\ \mathrm{dom}\ \tau\big] \text{ be a closed subspace of } E.$$

Then, for all $(x^*,y^*)\in E^*\times F^*$,

$$\rho^*(x^*,y^*)=\min\big\{\sigma^*(s^*,y^*)+\tau^*(t^*,y^*)\colon\ s^*,\,t^*\in E^*,\ s^*+t^*=x^*\big\}.$$

Proof. We first note that it is easy to see that ρ is convex. Furthermore, the conditions imply that $\mathrm{pr}_1\ \mathrm{dom}\ \sigma\cap\mathrm{pr}_1\ \mathrm{dom}\ \tau\neq\emptyset$, and so ρ is proper. Let $(x^*,y^*)\in E^*\times F^*$. We leave to the reader the simple verification that

$$\rho^*(x^*,y^*)\leq\inf\big\{\sigma^*(s^*,y^*)+\tau^*(t^*,y^*)\colon\ s^*,\,t^*\in E^*,\ s^*+t^*=x^*\big\}.$$

So what we have to prove is that there exists $t^*\in E^*$ such that

$$\sigma^*(x^*-t^*,y^*)+\tau^*(t^*,y^*)\leq\rho^*(x^*,y^*), \tag{4.2.1}$$

Since ρ is proper, $\rho^*(x^*,y^*)>-\infty$, so we can suppose that $\rho^*(x^*,y^*)\in\mathbb{R}$. Define $f,\ g\colon\ E\times F\times F\mapsto(-\infty,\infty]$ by

$$f(s,u,v):=\rho^*(x^*,y^*)-\langle s,x^*\rangle-\langle u+v,y^*\rangle+\sigma(s,u)$$

and

$$g(s,u,v):=\tau(s,v).$$

We note then that

$$\mathrm{dom}\ f=\{(s,u,v)\colon\ (s,u)\in\mathrm{dom}\ \sigma\}\quad\text{and}\quad\mathrm{dom}\ g=\{(s,u,v)\colon\ (s,v)\in\mathrm{dom}\ \tau\}.$$

We next prove that

$$\bigcup\nolimits_{\lambda>0}\lambda\big[\,\mathrm{dom}\ f-\mathrm{dom}\ g\big]=L\times F\times F, \tag{4.2.2}$$

which is a closed subspace of $E\times F\times F$. Since the inclusion "$\subset$" is immediate, it remains to prove "$\supset$". To this end, let $(s,u,v)\in L\times F\times F$. The definition of L gives $\lambda>0$, $(s_1,u_1)\in\mathrm{dom}\ \sigma$ and $(t_1,v_1)\in\mathrm{dom}\ \tau$ such that $s=\lambda(s_1-t_1)$. Let $u_2:=u_1-u/\lambda$ and $v_2:=v_1+v/\lambda$. Then $(s,u,v)=\lambda\big[(s_1,u_1,v_2)-(t_1,u_2,v_1)\big]\in\lambda\big[\,\mathrm{dom}\ f-\mathrm{dom}\ g\big]$, which completes the proof of (4.2.2). Now let $(s,u,v)\in E\times F\times F$. Then

$$(f+g)(s,u,v)=\rho^*(x^*,y^*)-\langle s,x^*\rangle-\langle u+v,y^*\rangle+\sigma(s,u)+\tau(s,v)$$

$$\geq \rho^*(x^*,y^*) - \langle s,x^*\rangle - \langle u+v,y^*\rangle + \rho(s,u+v) \geq 0.$$

Theorem 4.1 now gives $(t^*,u^*,v^*) \in E^* \times F^* \times F^*$ such that

$$f^*(-t^*,-u^*,-v^*) + g^*(t^*,u^*,v^*) \leq 0. \tag{4.2.3}$$

Since this implies that $f^*(-t^*,-u^*,-v^*) < \infty$, we must have

$$v^* = y^* \quad \text{and} \quad f^*(-t^*,-u^*,-v^*) = \sigma^*(x^*-t^*,y^*-u^*) - \rho^*(x^*,y^*).$$

(4.2.3) also implies that $g^*(t^*,u^*,v^*) < \infty$, from which

$$u^* = 0 \quad \text{and} \quad g^*(t^*,u^*,v^*) = \tau^*(t^*,v^*).$$

Thus (4.2.3) reduces to

$$\sigma^*(x^*-t^*,y^*-0) - \rho^*(x^*,y^*) + \tau^*(t^*,y^*) \leq 0.$$

This gives (4.2.1), and completes the proof of Theorem 4.2. □

Remark 4.3. We noted in the comments preceding Theorem 4.2 that Theorem 4.2 is, in fact, a generalization of Theorem 4.1. To see this, suppose that f, g and K are as in the statement of Theorem 4.1. Then we can obtain the result of Theorem 4.1 by applying Theorem 4.2 with $E = K$, $F = \{0\}$, for all $x \in E$, $\sigma(x,0) = f(x)$ and $\tau(x,0) = g(x)$, and $(x^*,y^*) = (0,0) \in E^* \times F^*$.

5. The maximal monotonicity of a sum in reflexive spaces

We start this section by using Fitzpatrick functions to obtain a sufficient condition for the sum of maximal monotone multifunctions on a reflexive space to be maximal monotone. However, the main result in this section is the "sandwiched closed subspace theorem", Theorem 5.5, a template for such existence theorems obtained by bootstrapping Lemma 5.1 through a sequence of four lemmas. Lemma 5.1 can also be established using a technique similar to that of [9, Theorem 15].

Lemma 5.1. *Let E be a reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ and $T\colon\ E \rightrightarrows E^*$ be maximal monotone and, writing $\mathrm{pr}_1(x,x^*) := x$,*

$$\textstyle\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \text{ dom } \phi_S - \mathrm{pr}_1 \text{ dom } \phi_T\big] \text{ be a closed subspace of } E. \tag{5.1.1}$$

Then $S+T$ is maximal monotone.

Proof. Let $\rho(x,x^*) := \inf\big\{\phi_S(x,s^*) + \phi_T(x,t^*)\colon\ s^*,\, t^* \in E^*,\ s^*+t^* = x^*\big\}$. From (3.0.1),

$$\rho(x,x^*) \geq \inf\big\{\langle x,s^*\rangle + \langle x,t^*\rangle\colon\ s^*,\, t^* \in E^*,\ s^*+t^* = x^*\big\} = \langle x,x^*\rangle.$$

We now derive from Theorem 4.2 and (3.0.5) that, for all $(x,x^*) \in E \times E^*$,

$$\begin{aligned}\rho^*(x^*,x) &= \min\big\{\phi_S^*(s^*,x) + \phi_T^*(t^*,x)\colon\ s^*,\, t^* \in E^*,\ s^*+t^* = x^*\big\}\\ &\geq \inf\big\{\langle x,s^*\rangle + \langle x,t^*\rangle\colon\ s^*,\, t^* \in E^*,\ s^*+t^* = x^*\big\} = \langle x,x^*\rangle.\end{aligned}$$

Theorem 1.4(a) with $k := \rho$ now gives that the set $\big\{(x,x^*) \in E \times E^*\colon\ \rho^*(x^*,x) = \langle x,x^*\rangle\big\}$ is maximal monotone. However, by direct computation from (3.0.6), this set is exactly $G(S+T)$, which completes the proof of Lemma 5.1. □

Lemma 5.2 is the first of the lemmas that we use to bootstrap Lemma 5.1 in our proof of Theorem 5.5, and is purely algebraic in character. In fact Lemma 5.2 is equivalent to the known fact that if C is convex then $a \in C$ and $b \in \operatorname{icr} C \Longrightarrow\ ]a,b] \subset \operatorname{icr} C$.

Lemma 5.2. *Let C be a convex subset of a vector space E, and $F := \bigcup_{\lambda>0} \lambda C$ be a subspace of E. Let $c \in C$ and $\alpha \in (0,1)$. Then*

$$\bigcup\nolimits_{\lambda>0} \lambda\big[C - \alpha c\big] = F. \tag{5.2.1}$$

Proof. $C - \alpha c \subset F - F = F$, which gives the inclusion "$\subset$" in (5.2.1). Now let $y \in F$. Then there exist $\mu > 0$ and $a \in C$ such that $y = \mu a$. Thus

$$(1-\alpha)a = \big[(1-\alpha)a + \alpha c\big] - \alpha c \in C - \alpha c$$

and so

$$y = \mu a \in \frac{\mu}{1-\alpha}\big[C - \alpha c\big] \subset \bigcup\nolimits_{\lambda>0} \lambda\big[C - \alpha c\big],$$

which gives the inclusion "$\supset$" in (5.2.1), and thus completes the proof of Lemma 5.2. □

Lemma 5.3 gives some connections between the sets used in Lemma 5.1. The technique used in Lemma 5.3(b) is taken from [13, Section 16, p. 57–62]. The technique used in Lemma 5.3(c) is taken from [13, Theorem 23.2, p. 87–88], which is not surprising given the identity "$\operatorname{pr}_1 \operatorname{dom} \phi_S = \operatorname{dom} \chi_S$" that we will establish in Remark 5.6.

Lemma 5.3. *Let E be a reflexive Banach space and $S\colon\ E \rightrightarrows E^*$ be maximal monotone. Then, writing $\operatorname{pr}_2(x^*, x) := x$, $D(S) := \operatorname{pr}_1 G(S) = \big\{x \in E\colon\ Sx \neq \emptyset\big\}$, "co" for "convex hull" and "lin" for "linear span":*

(a) $D(S) \subset \operatorname{co} D(S) \subset \operatorname{pr}_2 \operatorname{dom} \phi_S^* \subset \operatorname{pr}_1 \operatorname{dom} \phi_S$.
(b) *If F is a closed subspace of E, $w \in E$ and $D(S) \subset F + w$ then* $\operatorname{pr}_1 \operatorname{dom} \phi_S \subset F + w$.
(c) *Let $T\colon\ E \rightrightarrows E^*$ also be maximal monotone. Then*

$$\bigcup\nolimits_{\lambda>0} \lambda\big[\operatorname{pr}_1 \operatorname{dom} \phi_S - \operatorname{pr}_1 \operatorname{dom} \phi_T\big] \subset \overline{\operatorname{lin}\,\big(D(S) - D(T)\big)}.$$

Proof. It is clear from (3.0.5) that $\operatorname{pr}_2 \operatorname{dom} \phi_S^* \subset \operatorname{pr}_1 \operatorname{dom} \phi_S$ so, since $\operatorname{pr}_2 \operatorname{dom} \phi_S^*$ is convex, in order to prove (a) it remains to show that $D(S) \subset \operatorname{pr}_2 \operatorname{dom} \phi_S^*$. To this end, let $x \in D(S)$. Then there exists $x^* \in Sx$. From (3.0.6), $\phi_S^*(x^*, x) \in \mathbb{R}$, and so $x \in \operatorname{pr}_2 \operatorname{dom} \phi_S^*$. This completes the proof of (a). In order to prove (b), we shall write $F^\perp$ for the subspace $\big\{y^* \in E^*\colon\ \langle F, y^*\rangle = \{0\}\big\}$ of E^*. Let x be an arbitrary element of $\operatorname{pr}_1 \operatorname{dom} \phi_S$ and $u \in D(S)$. Then there exists $u^* \in Su$, and there also exists $x^* \in E^*$ such that $\phi_S(x, x^*) < \infty$. Let y^* be an arbitrary element of $F^\perp$. We first prove that

$$u^* + y^* \in Su. \tag{5.3.1}$$

To this end, let (s, s^*) be an arbitrary element of $G(S)$. Then, since

$$u - s \in D(S) - D(S) \subset F + w - (F + w) = F - F = F \quad \text{and} \quad y^* \in F^\perp,$$

we must have $\langle u-s, y^*\rangle = 0$ and so, since $u^* \in Su$, $s^* \in Ss$ and S is monotone,

$$\langle u-s, (u^*+y^*)-s^*\rangle = \langle u-s, u^*-s^*\rangle \geq 0.$$

The maximality of S now gives (5.3.1). We now derive from the definition of $\phi_S(x,x^*)$ that

$$\infty > \phi_S(x,x^*) \geq \langle u, x^*\rangle + \langle x, u^*+y^*\rangle - \langle u, u^*+y^*\rangle$$

from which

$$\infty > \phi_S(x,x^*) - \langle u, x^*\rangle - \langle x, u^*\rangle + \langle u, u^*\rangle \geq \langle x-u, y^*\rangle.$$

Since $F^\perp$ is a subspace of E^*, it follows that $\langle x-u, F^\perp\rangle = \{0\}$, and the bipolar theorem now implies that $x - u \in F$. Thus

$$x = (x-u) + u \in F + D(S) \subset F + (F+w) = (F+F) + w = F + w. \tag{5.3.2}$$

(b) now follows since (5.3.2) holds for all $x \in \mathrm{pr}_1\ \mathrm{dom}\ \phi_S$. For (c), we write F for the closed linear subspace $\overline{\mathrm{lin}\left(D(S)-D(T)\right)}$ of E. Let x be an arbitrary element of $\mathrm{pr}_1\ \mathrm{dom}\ \phi_S$ and y be an arbitrary element of $\mathrm{pr}_1\ \mathrm{dom}\ \phi_T$. Let t be an arbitrary element of $D(T)$. Then $D(S) - t \in D(S) - D(T) \subset F$. Consequently, $D(S) \subset F + t$, and it follows from (b) that $x \in F + t$, and so $t \in F + x$. Since t is an arbitrary element of $D(T)$, we have in fact proved that $D(T) \subset F + x$, and it follows from (b) (again) that $y \in F + x$, and so $x - y \in F$. Since this holds for all $x \in \mathrm{pr}_1\ \mathrm{dom}\ \phi_S$ and $y \in \mathrm{pr}_1\ \mathrm{dom}\ \phi_T$, we have established that $\mathrm{pr}_1\ \mathrm{dom}\ \phi_S - \mathrm{pr}_1\ \mathrm{dom}\ \phi_T \subset \overline{\mathrm{lin}\left(D(S)-D(T)\right)}$, from which (c) follows immediately. □

Lemma 5.4 explores how the concepts introduced in Lemma 5.1 react under a translation in the domain space.

Lemma 5.4. *Let E be a reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ be maximal monotone and $w \in E$. Define the maximal monotone multifunction $U\colon\ E \rightrightarrows E^*$ by*

$$(u,u^*) \in G(U) \iff (u+w, u^*) \in G(S).$$

Then:

(a) $(x,x^*) \in E \times E^* \implies \phi_U(x,x^*) = \phi_S(x+w,x^*) - \langle w, x^*\rangle.$
(b) $\mathrm{pr}_1\ \mathrm{dom}\ \phi_U = \mathrm{pr}_1\ \mathrm{dom}\ \phi_S - w.$
(c) $D(U) = D(S) - w.$

Proof. (a) Let $(x,x^*) \in E \times E^*$. Then

$$\begin{aligned}\phi_U(x,x^*) &= \sup\nolimits_{(u,u^*)\in G(U)} \left[\langle u, x^*\rangle + \langle x, u^*\rangle - \langle u, u^*\rangle\right]\\ &= \sup\nolimits_{(s,s^*)\in G(S)} \left[\langle s-w, x^*\rangle + \langle x, s^*\rangle - \langle s-w, s^*\rangle\right]\\ &= \sup\nolimits_{(s,s^*)\in G(S)} \left[\langle s, x^*\rangle + \langle x+w, s^*\rangle - \langle s, s^*\rangle\right] - \langle w, x^*\rangle\\ &= \phi_S(x+w,x^*) - \langle w, x^*\rangle.\end{aligned}$$

(b) follows from (a), and (c) is immediate from the definition of U. □

We now come to the main result of this section, the "sandwiched closed subspace theorem". We shall show in Remark 5.6 how different choices for F lead to known sufficient conditions for $S+T$ to be maximal monotone.

Theorem 5.5. *Let E be a reflexive Banach space, and $S\colon E \rightrightarrows E^*$ and $T\colon E \rightrightarrows E^*$ be maximal monotone. Suppose there exists a closed subspace F of E such that*

$$D(S) - D(T) \subset F \subset \bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big]. \tag{5.5.1}$$

Then $S+T$ is maximal monotone. Furthermore, for all $\varepsilon > 0$,

$$D(S) - D(T) \subset \mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T \subset (1+\varepsilon)\big[D(S) - D(T)\big], \tag{5.5.2}$$

(that is to say, $\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T$ *and* $D(S) - D(T)$ *are almost identical) and*

$$\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] = \bigcup_{\lambda>0} \lambda\big[D(S) - D(T)\big]. \tag{5.5.3}$$

Proof. (5.5.1) gives $\overline{\mathrm{lin}\,\big(D(S) - D(T)\big)} \subset F$. We then obtain from Lemma 5.3(c) that $\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] \subset F$, and another application of (5.5.1) implies that

$$\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] = F, \tag{5.5.4}$$

so (5.1.1) is satisfied, and the maximal monotonicity of $S+T$ follows from Lemma 5.1. Let $\varepsilon > 0$ and $\alpha := 1/(1+\varepsilon) \in (0,1)$. Let $c \in C := \mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T$. We now apply Lemma 5.2 and obtain from (5.5.4) that

$$\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T - \alpha c\big] = \bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] = F.$$

Define U as in Lemma 5.4, with $w := \alpha c$. Lemma 5.4(b) now gives that

$$\bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_U - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] = \bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \alpha c - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T\big] = F,$$

and so Lemma 5.1 (with S replaced by U) implies that $U+T$ is maximal monotone and, in particular, $D(U) \cap D(T) \neq \emptyset$. Using Lemma 5.4(c), we derive that

$$\big(D(S) - \alpha c\big) \cap D(T) \neq \emptyset,$$

from which $\alpha c \in D(S) - D(T)$, that is to say $c \in (1+\varepsilon)\big(D(S) - D(T)\big)$. Since this holds for any $c \in \mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T$, we have proved that

$$\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1 \ \mathrm{dom}\ \phi_T \subset (1+\varepsilon)\big(D(S) - D(T)\big).$$

(5.5.2) now follows from Lemma 5.3(a), and (5.5.3) is an immediate consequence of (5.5.2). □

Remark 5.6. We end up with some comments about possible choices for F in Theorem 5.5, recalling from Lemma 5.3(a) that $D(S) \subset \mathrm{co}\ D(S) \subset \mathrm{pr}_2 \ \mathrm{dom}\ \phi_S^* \subset \mathrm{pr}_1 \ \mathrm{dom}\ \phi_S$. If we take $F = \bigcup_{\lambda>0} \lambda\big[D(S) - D(T)\big]$, we obtain [13, (23.2.2), p. 87], while the choice $F = \bigcup_{\lambda>0} \lambda\big[\,\mathrm{co}\ D(S) - \mathrm{co}\ D(T)\big]$ gives us [13, (23.2.4), p. 87]. Either of these cases can be used to establish [9, Theorem 15] (but without the need to renorm E). We now discuss the choice $F = \bigcup_{\lambda>0} \lambda\big[\,\mathrm{pr}_2 \ \mathrm{dom}\ \phi_S^* - \mathrm{pr}_2 \ \mathrm{dom}\ \phi_T^*\big]$. Here, we define the function $c_S\colon E \times E^* \mapsto (-\infty, \infty]$ by

$$c_S(x,x^*) := \begin{cases} \langle x, x^*\rangle, & \text{if } (x,x^*) \in G(S); \\ \infty, & \text{otherwise.} \end{cases}$$

Then (see [6, Proposition 4.1, p. 63]) $\phi_S^*(x^*,x) = c_S^{**}(x,x^*) = \overline{\mathrm{co}}\, c_S(x,x^*)$ (in the notation of [17]) and so we obtain [17, Corollary 4]. Lemma 5.3(a) also leads us to the choice $F = \bigcup_{\lambda>0}\lambda\big[\,\mathrm{pr}_1 \ \mathrm{dom}\ \phi_S - \mathrm{pr}_1\ \mathrm{dom}\ \phi_T\big]$. In order to examine this, we must discuss briefly the technique of the *big convexification.* It was shown in [13, Section 9] how to define a convex subset C of a vector space, $\delta\colon\ G(S)\mapsto C$, affine maps $p\colon\ C\mapsto E$, $q\colon\ C\mapsto E^*$ and $r\colon\ C\mapsto\mathbb{R}$ such that

(5.6.1) $$C = \mathrm{co}\ \delta\big(G(S)\big)$$

and

(5.6.2)
$$(s,s^*)\in G(S)\quad\Longrightarrow\quad p\circ\delta(s,s^*) = s,\ \ q\circ\delta(s,s^*) = s^* \text{ and } r\circ\delta(s,s^*) = \langle s,s^*\rangle.$$

Now $x\in\mathrm{pr}_1\ \mathrm{dom}\ \phi_S$ if, and only if there exists $x^*\in E^*$ such that $\phi_S(x,x^*)<\infty$, or equivalently, such that, for some $M\ge 0$,

$$(s,s^*)\in G(S)\quad\Longrightarrow\quad \langle s,x^*\rangle + \langle x,s^*\rangle - \langle s,s^*\rangle \le M.$$

Using (5.6.2), this can be rewritten

$$(s,s^*)\in G(S)\quad\Longrightarrow\quad \big\langle p\circ\delta(s,s^*),x^*\big\rangle + \big\langle x,q\circ\delta(s,s^*)\big\rangle - r\circ\delta(s,s^*)\le M.$$

(5.6.1) implies that this is equivalent to

$$c\in C\quad\Longrightarrow\quad \big\langle -p(c),x^*\big\rangle - \big\langle x,q(c)\big\rangle + r(c)\ge -M.$$

Since the maps $c\mapsto -p(c)$ and $c\mapsto\big\langle x,q(c)\big\rangle - r(c)$ are affine, it follows from a special case of the new version of the Hahn–Banach theorem proved in [14, Theorem 1.5] (*if E is a nontrivial vector space, $S\colon\ E\mapsto\mathbb{R}$ is sublinear, C is a nonempty convex subset of a vector space, $g\colon\ C\mapsto E$ is affine and $f\colon\ C\mapsto\mathbb{R}$ is convex then there exists a linear functional L on E such that $L\le S$ on E and $\inf_C\big[f+L\circ g\big] = \inf_C\big[f+S\circ g\big]$*) that this is, in turn, equivalent to

there exists $N\ge 0$ such that $\quad c\in C\quad\Longrightarrow\quad N\|-p(c)\| - \big\langle x,q(c)\big\rangle + r(c)\ge -M.$

Combining together M and N into a single constant, we derive that $x\in\mathrm{pr}_1\ \mathrm{dom}\ \phi_S$ if, and only if

there exists $K\ge 0$ such that $\quad c\in C\quad\Longrightarrow\quad K + K\|p(c)\|\ge\big\langle x,q(c)\big\rangle - r(c),$

that is to say,

$$\sup_{c\in C}\frac{\big\langle x,q(c)\big\rangle - r(c)}{1+\|p(c)\|}<\infty,\quad\text{in other words,}\quad \chi_S(x)<\infty,$$

where the convex function χ_S is as defined in [13, Definition 15.1, p. 53]. So we have proved that $\mathrm{pr}_1\ \mathrm{dom}\ \phi_S = \mathrm{dom}\ \chi_S$, and so this choice of F gives us [13, (23.2.6), p. 87]. Of course, there are also valid "hybrid" choices, such as $F = \bigcup_{\lambda>0}\lambda\big[D(S) - \mathrm{pr}_1\ \mathrm{dom}\ \phi_T\big]$.

In all these cases, it follows from (5.5.4) that F is the same set, independently of how F is initially defined.

6. The Brezis–Crandall–Pazy condition

In this section, we investigate sufficient conditions for $S+T$ to be maximal monotone of a kind different from those considered in previous sections. The most general result in this section is Theorem 6.2, which is a generalization of [13, Theorem 24.3, p. 94]. We show in Corollary 6.5 how to deduce from this the result of Brezis, Crandall and Pazy, which has found applications to partial differential equations. We refer the reader to their original paper, [2], for more details.

Lemma 6.1. *Let E be a nontrivial reflexive Banach space, $U\colon\ E \rightrightarrows E^*$ and $V\colon\ E \rightrightarrows E^*$ be maximal monotone and $\operatorname{pr}_1 \operatorname{dom}\phi_U \cap \operatorname{pr}_1 \operatorname{dom}\phi_V \neq \emptyset$. Then there exists $R \geq 0$ (independent of n) with the following property: for all $n \geq 1$, there exist $(z_n, \xi_n) \in E \times E$ and $(z_n^*, \xi_n^*) \in E^* \times E^*$ such that*

$$\|z_n\|^2 + n^2\|\xi_n\|^2 + \|z_n^*\|^2 \leq R^2, \tag{6.1.1}$$

and

$$\psi_U(z_n, z_n^* - \xi_n^*) + \psi_V(z_n + \xi_n, \xi_n^*) + \Delta(z_n, z_n^*) + \Delta(n\xi_n, \xi_n^*/n) = 0. \tag{6.1.2}$$

Proof. Since $\operatorname{pr}_1 \operatorname{dom}\phi_U \cap \operatorname{pr}_1 \operatorname{dom}\phi_V \neq \emptyset$, we can choose $(u_0, u_0^*) \in \operatorname{dom}\phi_U$ and $(v_0, v_0^*) \in \operatorname{dom}\phi_V$ such that that $u_0 = v_0$. Let $Q := \sqrt{\|u_0\|^2 + \|u_0^* + v_0^*\|^2 + \|v_0^*\|^2}$. Q is clearly independent of n. Define $S_n\colon\ E \times E \rightrightarrows E^* \times E^*$ by

$$G(S_n) = \big\{\big((s, \sigma), (s^*, \sigma^*)\big)\colon\ (s, s^* - n\sigma^*) \in G(U),\ (s + \sigma/n, n\sigma^*) \in G(V)\big\}.$$

Using the equality $\big\langle(z, z + \zeta/n), (z^* - n\zeta^*, n\zeta^*)\big\rangle = \langle z, z^*\rangle + \langle\zeta, \zeta^*\rangle$, which is valid for all $\big((z, \zeta), (z^*, \zeta^*)\big) \in (E \times E) \times (E^* \times E^*)$, it is easy to check that S_n is maximal monotone and, for all $\big((z, \zeta), (z^*, \zeta^*)\big) \in (E \times E) \times (E^* \times E^*)$,

$$\phi_{S_n}\big((z, \zeta), (z^*, \zeta^*)\big) = \phi_U(z, z^* - n\zeta^*) + \phi_V(z + \zeta/n, n\zeta^*) \tag{6.1.3}$$

and

$$\psi_{S_n}\big((z, \zeta), (z^*, \zeta^*)\big) = \psi_U(z, z^* - n\zeta^*) + \psi_V(z + \zeta/n, n\zeta^*). \tag{6.1.4}$$

Let $\eta_n = \big((u_0, 0), (u_0^* + v_0^*, v_0^*/n)\big) \in (E \times E) \times (E^* \times E^*)$. Then

$$\|\eta_n\| = \sqrt{\|u_0\|^2 + \|u_0^* + v_0^*\|^2 + \|v_0^*\|^2/n^2} \leq Q,$$

so, even though η_n depends on n, $\{\|\eta_n\|\}_{n\geq 1}$ is bounded. Furthermore, (6.1.3) gives $\phi_{S_n}(\eta_n) = \phi_U(u_0, u_0^*) + \phi_V(v_0, v_0^*)$, which is independent of n. Then, from Theorem 3.1(a), there exists $\big((z_n, \zeta_n), (z_n^*, \zeta_n^*)\big) \in (E \times E) \times (E^* \times E^*)$ such that

$$\big\|\big((z_n, \zeta_n), (z_n^*, \zeta_n^*)\big)\big\| \leq \tfrac{1}{\sqrt{2}}\|\eta_n\| + \sqrt{\phi_{S_n}(\eta_n) + \tfrac{1}{2}\|\eta_n\|^2} \leq \tfrac{1}{\sqrt{2}}Q + \sqrt{\phi_{S_n}(\eta_n) + \tfrac{1}{2}Q^2}, \tag{6.1.5}$$

which is independent of n, and

$$\psi_{S_n}\big((z_n, \zeta_n), (z_n^*, \zeta_n^*)\big) + \Delta\big((z_n, \zeta_n), (z_n^*, \zeta_n^*)\big) = 0. \tag{6.1.6}$$

Let $\xi_n := \zeta_n/n$ and $\xi_n^* := n\zeta_n^*$. (6.1.1) follows by expanding out the terms in (6.1.5), and (6.1.2) follows by expanding out the terms in (6.1.6) and using (6.1.4). □

By saying that j is *increasing* in the statement of Theorem 6.2 below, we mean that

$$0 \le \rho_1 \le \rho_2,\ 0 \le \sigma_1 \le \sigma_2 \text{ and } 0 \le \tau_1 \le \tau_2 \quad \Longrightarrow \quad j(\rho_1, \sigma_1, \tau_1) \le j(\rho_2, \sigma_2, \tau_2).$$

Theorem 6.2. *Let E be a nontrivial reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ and $T\colon\ E \rightrightarrows E^*$ be maximal monotone and $\operatorname{pr}_1 \operatorname{dom} \phi_S \cap \operatorname{pr}_1 \operatorname{dom} \phi_T \ne \emptyset$. Suppose that there exists an increasing function $j\colon\ [0,\infty) \times [0,\infty) \times [0,\infty) \to [0,\infty)$ such that*

$$\text{(6.2.1)} \quad \begin{cases} (x, x^* - \xi^*) \in G(S),\ (x + \xi, \xi^*) \in G(T),\ \xi \ne 0 \text{ and } \langle \xi, \xi^* \rangle = -\|\xi\|\|\xi^*\| \\ \qquad\qquad \Longrightarrow \quad \|\xi^*\| \le j\big(\|x\|, \|x^*\|, \|\xi\|\|\xi^*\|\big). \end{cases}$$

Then $S + T$ is maximal monotone.

Proof. We will first prove that, for all $(w, w^*) \in E \times E^*$, there exists $(x, x^*, \xi^*) \in E \times E^* \times E^*$ such that (see (1.3.3) for the definition of $\delta_{(w,w^*)}$)

$$\text{(6.2.2)} \qquad \phi_S(x, x^* - \xi^*) + \phi_T(x, \xi^*) + \delta_{(w,w^*)}(x, x^*) \le 0.$$

So let (w, w^*) be an arbitrary element of $E \times E^*$. Define the maximal monotone multifunctions $U\colon\ E \rightrightarrows E^*$ and $V\colon\ E \rightrightarrows E^*$ by $G(U) := G(S) - (w, w^*)$ and $G(V) := G(T) - (w, 0)$. From a slight extension of the argument of Lemma 5.4(b),

$$\operatorname{pr}_1 \operatorname{dom} \phi_U \cap \operatorname{pr}_1 \operatorname{dom} \phi_V = (\operatorname{pr}_1 \operatorname{dom} \phi_S - w) \cap (\operatorname{pr}_1 \operatorname{dom} \phi_T - w) \ne \emptyset.$$

Let R be as in Lemma 6.1. From Lemma 6.1, for all $n \ge 1$, there exist $(z_n, \xi_n) \in E \times E$ and $(z_n^*, \xi_n^*) \in E^* \times E^*$ such that (6.1.1) and (6.1.2) are satisfied. For all $n \ge 1$, let $x_n = w + z_n$ and $x_n^* = w^* + z_n^*$. Then (6.1.2) becomes

$$\text{(6.2.3)} \quad \psi_S(x_n, x_n^* - \xi_n^*) + \psi_T(x_n + \xi_n, \xi_n^*) + \Delta(x_n - w, x_n^* - w^*) + \Delta(n\xi_n, \xi_n^*/n) = 0.$$

This implies that $\Delta(n\xi_n, \xi_n^*/n) = 0$ and so, from (1.1.2),

$$\text{(6.2.4)} \qquad \langle \xi_n, \xi_n^* \rangle = -\|\xi_n\|\|\xi_n^*\| \le 0 \quad \text{and} \quad \|\xi_n\|\|\xi_n^*\| = \|n\xi_n\|^2.$$

(6.2.3) also implies that $\psi_S(x_n, x_n^* - \xi_n^*) + \psi_T(x_n + \xi_n, \xi_n^*) + \Delta(x_n - w, x_n^* - w^*) = 0$, i.e., $\phi_S(x_n, x_n^* - \xi_n^*) + \phi_T(x_n + \xi_n, \xi_n^*) + \delta_{(w,w^*)}(x_n, x_n^*) - \langle \xi_n, \xi_n^* \rangle = 0$. Taking (6.2.4) into account, we derive that

$$\text{(6.2.5)} \qquad \phi_S(x_n, x_n^* - \xi_n^*) + \phi_T(x_n + \xi_n, \xi_n^*) + \delta_{(w,w^*)}(x_n, x_n^*) \le 0.$$

If there exists $n \ge 1$ such that $\xi_n = 0$ then this gives (6.2.2) with $(x, x^*, \xi^*) := (x_n, x_n^*, \xi_n^*)$. So we can and will assume that, for all $n \ge 1$, $\xi_n \ne 0$. It is clear from (6.1.1) that $\sup_{n\ge1} \|x_n\| \le R + \|w\|$, $\sup_{n\ge1} \|x_n^*\| \le R + \|w^*\|$ and $\sup_{n\ge1} \|n\xi_n\| \le R$. Using (6.2.4) and applying (6.2.1) gives

$$\sup_{n\ge1} \|\xi_n^*\| \le j\big(R + \|w\|, R + \|w^*\|, R^2\big) < \infty.$$

Thus, by passing to a subnet, we can suppose that $x_\alpha \rightharpoonup x$, $\xi_\alpha \to 0$, $x_\alpha^* \rightharpoonup x^*$ and $\xi_\alpha^* \rightharpoonup \xi^*$. For all α, (6.2.5) gives

$$\phi_S(x_\alpha, x_\alpha^* - \xi_\alpha^*) + \phi_T(x_\alpha + \xi_\alpha, \xi_\alpha^*) + \delta_{(w,w^*)}(x_\alpha, x_\alpha^*) \le 0.$$

We now obtain (6.2.2) by passing to the limit, and using (1.3.4) and the weak lower semicontinuity of ϕ_S and ϕ_T. Combining (1.1.1), (1.3.3), (3.0.1) and (6.2.2) gives us that

$$0 \le \langle x, x^* - \xi^* \rangle + \langle x, \xi^* \rangle + \delta_{(w,w^*)}(x, x^*)$$

$$\leq \phi_S(x, x^* - \xi^*) + \phi_T(x, \xi^*) + \delta_{(w,w^*)}(x, x^*) \leq 0.$$

Thus $\phi_S(x, x^* - \xi^*) = \langle x, x^* - \xi^* \rangle$ and $\phi_T(x, \xi^*) = \langle x, \xi^* \rangle$, and (3.0.2) implies that $(x, x^* - \xi^*) \in G(S)$ and $(x, \xi^*) \in G(T)$, from which $(x, x^*) \in G(S + T)$. Define the convex function $h\colon\ E \times E^* \mapsto (-\infty, \infty]$ by

$$h(x, x^*) := \inf \left\{ \phi_S(x, x^* - v^*) + \phi_T(x, v^*)\colon\ v^* \in E^* \right\} \geq \langle x, x^* \rangle.$$

(h is identical with the function ρ of Lemma 5.1. It is clear from (3.0.2) that $G(S + T) \subset M_h$, and (6.2.2) implies (1.2.2) since $h(x, x^*) \leq \phi_S(x, x^* - \xi^*) + \phi_T(x, \xi^*)$. Lemma 1.2(c) with $G := G(S + T)$ now gives that $G(S + T)$ is maximal monotone. □

Remark 6.3. We note that (6.2.1) is satisfied if we assume that the first line of (6.2.1) implies that $\|\xi^*\|$ is bounded by certain special functions of $\|\xi\|$ only. Let a and b be large positive numbers, $\lambda, \mu \geq 0$ and $\mu \leq a \vee \lambda^b$. Then

$$\mu > a \quad \Longrightarrow \quad \mu \leq \lambda^b \quad \Longrightarrow \quad \mu^{\frac{1}{b+1}} \leq \lambda^{\frac{b}{b+1}} \quad \Longrightarrow \quad \mu \leq (\lambda\mu)^{\frac{b}{b+1}}.$$

Consequently, $\mu \leq a \vee (\lambda\mu)^{\frac{b}{b+1}}$. Thus, if the first line of (6.2.1) implies that $\|\xi^*\| \leq a \vee \|\xi\|^b$, then (6.2.1) is satisfied with $j(\cdot, \cdot, \theta) := a \vee \theta^{\frac{b}{b+1}}$.

In what follows, if $U\colon\ E \rightrightarrows E^*$ and $x \in E$, we write $|Ux| = \inf \|Ux\|$. The next result is an implicit version of the Brezis–Crandall–Pazy theorem on the perturbation of multifunctions ("implicit" because the quantity $|Tx|$ appears on both sides of the inequality in (6.4.1)). The original explicit version will appear in Corollary 6.5, and a new explicit version in Corollary 6.6.

Corollary 6.4. *Let E be a nontrivial reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ and $T\colon\ E \rightrightarrows E^*$ be maximal monotone, $D(S) \subset D(T)$, and suppose that there exists an increasing function $j\colon\ [0, \infty) \times [0, \infty) \to [0, \infty)$ such that,*

$$x \in D(S) \quad \Longrightarrow \quad |Tx| \leq j\big(\|x\|, (|Sx| - |Tx|)^+\big). \tag{6.4.1}$$

Then $S + T$ is maximal monotone.

Proof. We first note from Lemma 5.3(a) that $\mathrm{pr}_1 \operatorname{dom} \phi_S \cap \mathrm{pr}_1 \operatorname{dom} \phi_T \supset D(S) \neq \emptyset$. We now show that (6.2.1) is satisfied. To this end, suppose that

$$(x, x^* - \xi^*) \in G(S),\ (x + \xi, \xi^*) \in G(T),\ \xi \neq 0 \text{ and } \langle \xi, \xi^* \rangle = -\|\xi\|\|\xi^*\|.$$

This clearly implies that $x \in D(S) \subset D(T)$. Now let t^* be an arbitrary element of Tx. We then have $(x, t^*) \in G(T)$. Since $(x + \xi, \xi^*) \in G(T)$ and T is monotone, $\langle \xi, \xi^* - t^* \rangle \geq 0$, and so $-\langle \xi, \xi^* \rangle \leq -\langle \xi, t^* \rangle$. Thus $\|\xi\|\|\xi^*\| \leq \|\xi\|\|t^*\|$, and division by $\|\xi\|$ gives $\|\xi^*\| \leq \|t^*\|$. If we now take the infimum over $t^* \in Tx$, we obtain $\|\xi^*\| \leq |Tx|$. Since $(x, x^* - \xi^*) \in G(S)$, we also have $|Sx| \leq \|x^* - \xi^*\| \leq \|x^*\| + \|\xi^*\| \leq \|x^*\| + |Tx|$, and so $|Sx| - |Tx| \leq \|x^*\|$, from which $(|Sx| - |Tx|)^+ \leq \|x^*\|$. Thus (6.4.1) implies that $\|\xi^*\| \leq |Tx| \leq j\big(\|x\|, \|x^*\|\big)$, and it now follows from Theorem 6.2 that $S + T$ is maximal monotone. □

Corollary 6.5. *Let E be a nontrivial reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ and $T\colon\ E \rightrightarrows E^*$ be maximal monotone, $D(S) \subset D(T)$, and suppose that there exist increasing functions $k\colon\ [0, \infty) \to [0, 1)$ and $C\colon\ [0, \infty) \to [0, \infty)$ such that,*

$$x \in D(S) \Longrightarrow |Tx| \leq k(\|x\|)|Sx| + C(\|x\|). \tag{6.5.1}$$

Then $S+T$ is maximal monotone.

Proof. Let $x \in D(S)$. From (6.5.1), $\big(1-k(\|x\|)\big)|Tx| \le k(\|x\|)\big(|Sx|-|Tx|\big)+C(\|x\|) \le k(\|x\|)(|Sx|-|Tx|)^+ + C(\|x\|)$, and the result now follows from Corollary 6.4 with

$$j(\rho,\sigma) := \frac{k(\rho)\sigma + C(\rho)}{1-k(\rho)}. \qquad \square$$

In our final result, we allow k to take values bigger than 1, but we replace $|Sx|$ by $|Sx|^p$ in the statement of Corollary 6.5.

Corollary 6.6. *Let E be a nontrivial reflexive Banach space, $S\colon\ E \rightrightarrows E^*$ and $T\colon\ E \rightrightarrows E^*$ be maximal monotone, $D(S) \subset D(T)$, and suppose that $0<p<1$ and there exist increasing functions $k\colon\ [0,\infty) \to [0,\infty)$ and $C\colon\ [0,\infty)\to[0,\infty)$ such that,*

$$x \in D(S) \Longrightarrow |Tx| \le k(\|x\|)|Sx|^p + C(\|x\|). \tag{6.6.1}$$

Then $S+T$ is maximal monotone.

Proof. Let $x\in D(S)$. From (6.6.1) and the fact that $\lambda,\mu \ge 0 \Longrightarrow (\lambda+\mu)^p \le \lambda^p+\mu^p$,

$$\begin{aligned}|Tx| &\le k(\|x\|)(|Tx|\vee|Sx|)^p + C(\|x\|) = k(\|x\|)\big(|Tx| + (|Sx|-|Tx|)^+\big)^p + C(\|x\|)\\ &\le k(\|x\|)|Tx|^p + k(\|x\|)\big((|Sx|-|Tx|)^+\big)^p + C(\|x\|).\end{aligned}$$

Now if $k(\|x\|)|Tx|^p \le \frac{1}{2}|Tx|$ then this gives $|Tx| \le 2k(\|x\|)\big((|Sx|-|Tx|)^+\big)^p + 2C(\|x\|)$, while if $\frac{1}{2}|Tx| < k(\|x\|)|Tx|^p$ then, of course, $|Tx| < \big(2k(\|x\|)\big)^{1/(1-p)}$. Thus the result follows from Corollary 6.4, with $j(\rho,\sigma) := \big[2k(\rho)\sigma^p + 2C(\rho)\big] \vee \big(2k(\rho)\big)^{1/(1-p)}$. $\square$

Remark 6.7. We emphasize that, unlike the analysis in [2], we do not use any renorming or fixed–point theorems in any of the above results. Theorem 6.2 does not have the limitation $D(S)\subset D(T)$ of Corollary 6.5, though we do not know if it has any practical applications other than those that can be obtained from Corollaries 6.5 and 6.6.

7. Other formulas for $\min\big\{\|x\|\colon\ x\in E,\ (S+J)x \ni 0\big\}$

Let E be a nontrivial reflexive Banach space and $S\colon\ E\rightrightarrows E^*$ be a maximal monotone multifunction. We showed in Theorem 3.1 that

$$\min\big\{\|x\|\colon\ x\in E,\ (S+J)x\ni 0\big\} = \tfrac{1}{\sqrt{2}}\sup\nolimits_{\eta\in E\times E^*}\Big[\|\eta\| - \sqrt{2\phi_S(\eta)+\|\eta\|^2}\Big]^+. \tag{7.0.1}$$

In this final section, we give a general result that leads to other formulas for the left–hand side, which might be more convenient for computation. In particular, we will see that if $\|(x,x^*)\|_1 := \|x\| + \|x^*\|$ and $\|(x,x^*)\|_\infty := \|x\| \vee \|x^*\|$ then

$$\min\big\{\|x\|\colon\ x \in E,\ (S+J)x \ni 0\big\}$$

$$= \tfrac{1}{2} \sup_{\eta \in E \times E^*} \left[\|\eta\|_1 - \sqrt{4\phi_S(\eta) + \|\eta\|_1{}^2}\right]^+ \tag{7.0.2}$$

$$= \sup_{\eta \in E \times E^*} \left[\|\eta\|_\infty - \sqrt{\phi_S(\eta) + \|\eta\|_\infty{}^2}\right]^+. \tag{7.0.3}$$

We start off by investigating some elementary properties of norms on $\mathbb{R}^2$. Let $\mathcal{N}$ be a norm on $\mathbb{R}^2$. We say that $\mathcal{N}$ is *octagonal* if

$$(\lambda_1, \lambda_2) \in \mathbb{R}^2 \quad\Longrightarrow\quad \mathcal{N}(\lambda_1, \lambda_2) = \mathcal{N}(\lambda_2, \lambda_1) = \mathcal{N}(|\lambda_1|, |\lambda_2|),$$

and we write $C_\mathcal{N} := \mathcal{N}(1,1)$. If $\mathcal{N}(\lambda_1, \lambda_2) = \sqrt{\lambda_1{}^2 + \lambda_2{}^2}$ then $C_\mathcal{N} = \sqrt{2}$, if $\mathcal{N}(\lambda_1, \lambda_2) = |\lambda_1| + |\lambda_2|$ then $C_\mathcal{N} = 2$, while if $\mathcal{N}(\lambda_1, \lambda_2) = |\lambda_1| \vee |\lambda_2|$ then $C_\mathcal{N} = 1$. If we substitute these three values of $\mathcal{N}$ in Theorem 7.3 below we obtain, respectively, (7.0.1), (7.0.2) and (7.0.3). If $\mathcal{N}$ is octagonal, $0 \le \lambda_1 \le \mu_1$ and $0 \le \lambda_2 \le \mu_2$ then (λ_1, λ_2) is a convex combination of (μ_1, μ_2), $(-\mu_1, \mu_2)$ and $(\mu_1, -\mu_2)$, consequently $\mathcal{N}(\lambda_1, \lambda_2) \le \mathcal{N}(\mu_1, \mu_2)$. In order to prove Theorem 7.3, we will need to discuss the dual norm $\mathcal{N}^*$ on $\mathbb{R}^2$, defined by

$$\mathcal{N}^*(\lambda_1^*, \lambda_2^*) := \max_{\mathcal{N}(\lambda_1, \lambda_2) \le 1} \lambda_1\lambda_1^* + \lambda_2\lambda_2^*.$$

If $\mathcal{N}$ is octagonal then

$$\mathcal{N}^*(\lambda_1^*, \lambda_2^*) = \max_{\mathcal{N}(|\lambda_1|, |\lambda_2|) \le 1} \lambda_1\lambda_1^* + \lambda_2\lambda_2^* = \max_{\mathcal{N}(|\lambda_1|, |\lambda_2|) \le 1} |\lambda_1||\lambda_1^*| + |\lambda_2||\lambda_2^*|,$$

from which it follows easily that $\mathcal{N}^*$ is octagonal.

Lemma 7.1. *Let $\mathcal{N}$ be a octagonal norm on $\mathbb{R}^2$. Then:*

(a) *For all $\lambda_1, \lambda_2 \ge 0$, $\quad \mathcal{N}(\lambda_1, \lambda_2) \ge \frac{1}{2}(\lambda_1 + \lambda_2)C_\mathcal{N}$.*

(b) *$C_\mathcal{N} C_{\mathcal{N}^*} = 2$. Let $\gamma_\mathcal{N} := C_\mathcal{N}/C_{\mathcal{N}^*}$: then $\frac{1}{2}C_\mathcal{N}{}^2 = \gamma_\mathcal{N}$.*

(c) *For all $\lambda_1, \lambda_2 \ge 0$,*

$$\tfrac{1}{2}\mathcal{N}(\lambda_1, \lambda_2)^2 \ge \gamma_\mathcal{N}\lambda_1\lambda_2, \tag{7.1.1}$$

with equality if, and only if, $\lambda_1 = \lambda_2$.

Proof. (a) Let $\lambda_1, \lambda_2 \ge 0$. Then

$$\begin{aligned}\mathcal{N}(\lambda_1, \lambda_2) &= \tfrac{1}{2}\mathcal{N}(\lambda_1, \lambda_2) + \tfrac{1}{2}\mathcal{N}(\lambda_2, \lambda_1)\\ &\ge \mathcal{N}\big(\tfrac{1}{2}(\lambda_1, \lambda_2) + \tfrac{1}{2}(\lambda_2, \lambda_1)\big) = \mathcal{N}\big(\tfrac{1}{2}(\lambda_1 + \lambda_2, \lambda_1 + \lambda_2)\big),\end{aligned}$$

which gives (a).

(b) From (a), for all $(\lambda_1, \lambda_2) \in \mathbb{R}^2$,

$$\begin{aligned}\big\langle(\lambda_1, \lambda_2), \tfrac{1}{2}(C_\mathcal{N}, C_\mathcal{N})\big\rangle &= \tfrac{1}{2}(\lambda_1 + \lambda_2)C_\mathcal{N}\\ &\le \tfrac{1}{2}(|\lambda_1| + |\lambda_2|)C_\mathcal{N} \le \mathcal{N}(|\lambda_1|, |\lambda_2|) = \mathcal{N}(\lambda_1, \lambda_2),\end{aligned}$$

thus $\mathcal{N}^*\big(\tfrac{1}{2}(C_\mathcal{N}, C_\mathcal{N})\big) \le 1$, which gives $C_\mathcal{N} C_{\mathcal{N}^*} \le 2$. On the other hand,

$$C_\mathcal{N} C_{\mathcal{N}^*} = \mathcal{N}(1,1)\mathcal{N}^*(1,1) \ge \big\langle(1,1),(1,1)\big\rangle = 2,$$

which completes the proof of the first equality, and the second follows from the definition of $\gamma_{\mathcal{N}}$.

(c) Since $(\lambda_1+\lambda_2)^2 \ge 4\lambda_1\lambda_2$, (7.1.1) is immediate from (a) and (b). If $\lambda_1 = \lambda_2$ then we obviously have equality in (7.1.1). If, conversely, we have equality in (7.1.1) then, from (a) again, $\gamma_{\mathcal{N}}\lambda_1\lambda_2 = \frac{1}{2}\mathcal{N}(\lambda_1,\lambda_2)^2 \ge \frac{1}{8}(\lambda_1+\lambda_2)^2 C_{\mathcal{N}}{}^2 = \frac{1}{4}\gamma_{\mathcal{N}}(\lambda_1+\lambda_2)^2$. Thus $4\lambda_1\lambda_2 \ge (\lambda_1+\lambda_2)^2$, which implies that $\lambda_1 = \lambda_2$. □

If $\mathcal{N}$ is a octagonal norm on $\mathbb{R}^2$, we define a norm $\|\cdot\|_{\mathcal{N}}$ on $E\times E^*$ by $\|(x,x^*)\|_{\mathcal{N}} := \mathcal{N}\big(\|x\|,\|x^*\|\big)$. Since $\|\cdot\|_{\mathcal{N}}$ and $\|\cdot\|$ are equivalent norms, $(E\times E^*)^* = E^*\times E$ as before. The next result tells us that the dual norm, $\|\cdot\|_{\mathcal{N}}^*$, of $\|\cdot\|_{\mathcal{N}}$ on $E^*\times E$ is exactly what we would like.

Lemma 7.2. *Let $\mathcal{N}$ be a octagonal norm on $\mathbb{R}^2$. Then, for all $(u^*,u)\in E^*\times E$, $\|(u^*,u)\|_{\mathcal{N}}^* = \mathcal{N}^*\big(\|u^*\|,\|u\|\big)$.*

Proof. We have

$$\begin{aligned}\|(u^*,u)\|_{\mathcal{N}}^* &:= \max_{\|(x,x^*)\|_{\mathcal{N}}\le 1} \langle x,u^*\rangle + \langle u,x^*\rangle = \max_{\mathcal{N}(\|x\|,\|x^*\|)\le 1} \langle x,u^*\rangle + \langle u,x^*\rangle\\ &\le \max_{\mathcal{N}(\|x\|,\|x^*\|)\le 1} \|x\|\|u^*\| + \|u\|\|x^*\|\\ &\le \max_{\mathcal{N}(\lambda_1,\lambda_2)\le 1} \lambda_1\|u^*\| + \lambda_2\|u\| = \mathcal{N}^*\big(\|u^*\|,\|u\|\big).\end{aligned}$$

On the other hand, it follows from the last equality above that there exists $(\lambda_1,\lambda_2)\in\mathbb{R}^2$ such that $\mathcal{N}(\lambda_1,\lambda_2)\le 1$ and $\mathcal{N}^*\big(\|u^*\|,\|u\|\big) = \lambda_1\|u^*\| + \lambda_2\|u\|$. Now we can choose $(x,x^*)\in E\times E^*$ such that $\|x\| = |\lambda_1|$, $\|x^*\| = |\lambda_2|$, $\langle x,u^*\rangle = \lambda_1\|u^*\|$ and $\langle u,x^*\rangle = \lambda_2\|u\|$. But then, since $\|(x,x^*)\|_{\mathcal{N}} = \mathcal{N}\big(\|x\|,\|x^*\|\big) = \mathcal{N}(|\lambda_1|,|\lambda_2|)\le 1$,

$$\mathcal{N}^*\big(\|u^*\|,\|u\|\big) = \langle x,u^*\rangle + \langle u,x^*\rangle = \big\langle (x,x^*),(u^*,u)\big\rangle \le \|(u^*,u)\|_{\mathcal{N}}^*,$$

which completes the proof of Lemma 7.2. □

In what follows, of course $\gamma_{\mathcal{N}^*} := C_{\mathcal{N}^*}/C_{\mathcal{N}} = 1/\gamma_{\mathcal{N}}$.

Theorem 7.3. *Let E be a nontrivial reflexive Banach space and $S\colon E \rightrightarrows E^*$ be a maximal monotone multifunction. Let $\mathcal{N}$ be any octagonal norm on $\mathbb{R}^2$ and*

$$P_{\mathcal{N}} := \frac{1}{C_{\mathcal{N}}}\sup_{\eta\in E\times E^*}\left[\|\eta\|_{\mathcal{N}} - \sqrt{C_{\mathcal{N}}{}^2\phi_S(\eta) + \|\eta\|_{\mathcal{N}}{}^2}\right]^+.$$

Then

$$\min\big\{\|x\|\colon\ x\in E,\ (S+J)x\ni 0\big\} = P_{\mathcal{N}},$$

and so $P_{\mathcal{N}}$ is independent of $\mathcal{N}$.

Proof. It follows from (3.0.1) and Lemma 7.1(c), with $(\lambda_1,\lambda_2) = (\|x\|,\|x^*\|)$, that

$$\begin{aligned}&(x,x^*)\in E\times E^*\\ &\quad\Longrightarrow \gamma_{\mathcal{N}}\phi_S(x,x^*) + \tfrac{1}{2}\|(x,x^*)\|_{\mathcal{N}}{}^2 \ge \gamma_{\mathcal{N}}\langle x,x^*\rangle + \tfrac{1}{2}\mathcal{N}(\|x\|,\|x^*\|)^2\\ &\qquad\qquad\ge \tfrac{1}{2}\mathcal{N}(\|x\|,\|x^*\|)^2 - \gamma_{\mathcal{N}}\|x\|\|x^*\| \ge 0.\end{aligned}$$

Thus (3.0.3), and Theorem 2.1(a) with $F := \big(E \times E^*, \|\cdot\|_{\mathcal{N}}\big)$ and $f := \gamma_{\mathcal{N}}\phi_S = \frac{1}{2}{C_{\mathcal{N}}}^2\phi_S$, give $\eta^* \in E^* \times E$ such that

$$\|\eta^*\|_{\mathcal{N}}^* \le \sup_{\eta \in E\times E^*}\left[\|\eta\|_{\mathcal{N}} - \sqrt{{C_{\mathcal{N}}}^2\phi_S(\eta) + {\|\eta\|_{\mathcal{N}}}^2}\right]^+ = C_{\mathcal{N}}P_{\mathcal{N}} \tag{7.3.1}$$

and $(\gamma_{\mathcal{N}}\phi_S)^*(\eta^*) + \frac{1}{2}{\|\eta^*\|_{\mathcal{N}}^*}^2 \le 0$. Writing $\zeta^* = \gamma_{\mathcal{N}^*}\eta^*$, or equivalently, $\eta^* = \gamma_{\mathcal{N}}\zeta^*$, this becomes $\gamma_{\mathcal{N}}\phi_S^*(\zeta^*) + \frac{1}{2}\gamma_{\mathcal{N}}^2{\|\zeta^*\|_{\mathcal{N}}^*}^2 \le 0$, that is to say, $\gamma_{\mathcal{N}^*}\phi_S^*(\zeta^*) + \frac{1}{2}{\|\zeta^*\|_{\mathcal{N}}^*}^2 \le 0$. Let $(z, z^*) \in E \times E^*$ be such that $\zeta^* = (z^*, z)$. Then, using Lemma 7.2, we derive that $\gamma_{\mathcal{N}^*}\phi_S^*(z^*, z) + \frac{1}{2}\mathcal{N}^*(\|z^*\|, \|z\|)^2 \le 0$. But since the left hand side of this inequality is

$$\gamma_{\mathcal{N}^*}\left[\phi_S^*(z^*, z) - \langle z, z^*\rangle\right] + \gamma_{\mathcal{N}^*}\left[\|z\|\|z^*\| + \langle z, z^*\rangle\right] + \left[\tfrac{1}{2}\mathcal{N}^*(\|z^*\|, \|z\|)^2 - \gamma_{\mathcal{N}^*}\|z\|\|z^*\|\right],$$

and, from (3.0.5) and Lemma 7.1(c), with $\mathcal{N}$ replaced by $\mathcal{N}^*$ and $(\lambda_1, \lambda_2) = (\|z^*\|, \|z\|)$, each of the three summands is nonnegative, it follows that

$$\phi_S^*(z^*, z) = \langle z, z^*\rangle, \quad \|z\|\|z^*\| = -\langle z, z^*\rangle, \quad \text{and} \quad \tfrac{1}{2}\mathcal{N}^*(\|z^*\|, \|z\|)^2 = \gamma_{\mathcal{N}^*}\|z^*\|\|z\|.$$

Taking into account (3.0.6) and Lemma 7.1(c), with $\mathcal{N}$ replaced by $\mathcal{N}^*$ and $(\lambda_1, \lambda_2) = (\|z^*\|, \|z\|)$ again, we have $(z, z^*) \in G(S)$, $\|z\|\|z^*\| = -\langle z, z^*\rangle$ and $\|z^*\| = \|z\|$, that is to say, $-z^* \in Jz$. Since $0 = z^* + (-z^*)$, it is now immediate that $(S + J)z \ni 0$. Further, $\|\zeta^*\|_{\mathcal{N}}^* = \mathcal{N}^*(\|z^*\|, \|z\|) = \mathcal{N}^*(\|z\|, \|z\|) = C_{\mathcal{N}^*}\|z\|$ and so, from (7.3.1),

$$\|z\| = \frac{1}{C_{\mathcal{N}^*}}\|\zeta^*\|_{\mathcal{N}}^* = \frac{\gamma_{\mathcal{N}^*}}{C_{\mathcal{N}^*}}\|\eta^*\|_{\mathcal{N}}^* = \frac{1}{C_{\mathcal{N}}}\|\eta^*\|_{\mathcal{N}}^* \le P_{\mathcal{N}}.$$

In order to complete the proof, we must show that

$$x \in E \text{ and } (S + J)x \ni 0 \quad \Longrightarrow \quad \|x\| \ge P_{\mathcal{N}}. \tag{7.3.2}$$

So suppose that $x \in E$ and $(S + J)x \ni 0$. Then there exists $x^* \in Sx$ such that $-x^* \in Jx$. Write $\eta^* = \gamma_{\mathcal{N}}(x^*, x)$. Then, from Lemma 7.2 and the fact that $\|x^*\| = \|x\|$,

$$\begin{aligned}\|\eta^*\|_{\mathcal{N}}^* = \gamma_{\mathcal{N}}\|(x^*, x)\|_{\mathcal{N}}^* &= \gamma_{\mathcal{N}}\mathcal{N}^*(\|x^*\|, \|x\|)\\ &= \gamma_{\mathcal{N}}\mathcal{N}^*(\|x\|, \|x\|) = \gamma_{\mathcal{N}}C_{\mathcal{N}^*}\|x\| = C_{\mathcal{N}}\|x\|.\end{aligned}$$

Consequently, from the fact that $\|x\|^2 = -\langle x, x^*\rangle$ and (3.0.6),

$$\tfrac{1}{2}{\|\eta^*\|_{\mathcal{N}}^*}^2 = \tfrac{1}{2}{C_{\mathcal{N}}}^2\|x\|^2 = \gamma_{\mathcal{N}}\|x\|^2 = -\gamma_{\mathcal{N}}\langle x, x^*\rangle = -\gamma_{\mathcal{N}}{\phi_S}^*(x^*, x) = -(\gamma_{\mathcal{N}}\phi_S)^*(\eta^*).$$

So we have proved that $(\gamma_{\mathcal{N}}\phi_S)^*(\eta^*) + \frac{1}{2}{\|\eta^*\|_{\mathcal{N}}^*}^2 = 0$. Theorem 2.1(b) now gives $\|\eta^*\|_{\mathcal{N}}^* \ge C_{\mathcal{N}}P_{\mathcal{N}}$, from which (7.3.2) follows, completing the proof of Theorem 7.3. □

References

[1] H. Attouch and H. Brezis, *Duality for the sum of convex funtions in general Banach spaces*, Aspects of Mathematics and its Applications, J. A. Barroso, ed., Elsevier Science Publishers (1986), 125–133.

[2] H. Brezis, M. G. Crandall and A. Pazy, *Perturbations of nonlinear maximal monotone sets in Banach spaces*, Comm. Pur. Appl. Math. **23** (1970), 123–144.

[3] H. Brezis and A. Haraux, *Image d'une somme d'opérateurs monotones at applications*, Israel J. Math. **23** (1976), 165–186.
[4] R. S. Burachik and B. F. Svaiter, *Maximal monotone operators, convex function and a special family of enlargements*, Set–Valued Anal. **10** (2002), 297–316.
[5] R. S. Burachik and B. F. Svaiter, *Maximal monotonicity, conjugation and the duality product*, IMPA Preprint 129/2002, February 28, 2002.
[6] S. Fitzpatrick, *Representing monotone operators by convex functions*, Workshop Miniconference on Functional Analysis and Optimization (Canberra, 1988), 59–65, Proc. Centre Math. Anal. Austral. Nat. Univ., **20**, Austral. Nat. Univ., Canberra, 1988.
[7] E. Krauss, *A representation of maximal monotone operators by saddle functions*, Rev. Roum. Math. Pures Appl. **30** (1985), 823–837.
[8] J-E. Martínez-Legaz and M. Théra, *A convex representation of maximal monotone operators*, J. Nonlinear Convex Anal. **2** (2001), 243–247.
[9] J.-P. Penot, *The relevance of convex analysis for the study of monotonicity*, preprint.
[10] S. Reich, *The range of sums of accretive and monotone operators*, J. Math. Anal. Appl. **68** (1979), 310–317.
[11] S. Reich and S. Simons, *Fenchel duality, Fitzpatrick functions and the Kirszbraun–Valentine extension theorem*, Proc. Amer. Math. Soc., in press.
[12] R. T. Rockafellar, *Extension of Fenchel's duality theorem for convex functions*, Duke Math. J. **33** (1966), 81–89.
[13] S. Simons, *Minimax and monotonicity*, Lecture Notes in Mathematics **1693** (1998), Springer–Verlag.
[14] S. Simons, *A new version of the Hahn–Banach theorem*, Arch. Math. **80** (2003), 630–646.
[15] S. Simons and C. Zălinescu, *A new proof for Rockafellar's characterization of maximal monotone operators*, Proc. Amer. Math. Soc, **132** (2004), 2969–2972.
[16] C. Zălinescu, *Convex analysis in general vector spaces*, (2002), World Scientific.
[17] C. Zălinescu, *A new proof of the maximal monotonicity of the sum using the Fitzpatrick function*, in Variational Analysis and Appls., F. Giannessi and A. Maugeri, Eds., Kluwer Acad. Publ. Dordrecht, 2004, accepted for publication.

Manuscript received March 2, 2004
revised May 27, 2004

Stephen Simons
Department of Mathematics, University of California, Santa Barbara, CA 93106, USA
E-mail address: simons@math.ucsb.edu

Constantin Zălinescu
University "Al. I. Cuza" Iaşi, Faculty of Mathematics, Bd. Carol I Nr. 11, 700506 Iaşi, Romania
E-mail address: zalinesc@uaic.ro