MAXIMALITY OF SUMS
OF TWO MAXIMAL MONOTONE OPERATORS
IN GENERAL BANACH SPACE

JONATHAN M. BOR WEIN, FRSC

(Communicated by N. Tomczak-Jaegermann)

Abstract. We combine methods from convex analysis, based on a function of Simon Fitzpatrick, with a fine recent idea due to Voisei, to prove maximality of the sum of two maximal monotone operators in Banach space under various natural domain and transversality conditions.

1. Introduction and preliminaries

The results of this paper, especially Theorem 9, marry recent work by Voisei [2] with additional convex analysis described in [1, 2], see also [3, §5.1] or [4, §8.3], to provide an accessible short proof of the maximality of the sum of two maximal monotone operators under domain conditions such as $D(B) \cap \text{core } D(A) \neq \emptyset$, while either $D(A) \cap D(B)$ is closed and convex or $\text{core conv } D(B) \neq \emptyset$.

Recall that the domain of an extended-valued convex function, $\text{dom}(f)$, is the set of points with value less than $+\infty$, and that a point $s$ is in the core of a set $S$ (denoted by $s \in \text{core } S$) provided that $s$ lies in $S$ and $X = \bigcup_{\lambda > 0} \lambda(S - s)$. Recall that $x^* \in X^*$ is a subgradient of $f : X \to (-\infty, +\infty]$ at $x \in \text{dom } f$ provided that $f(y) - f(x) \geq \langle x^*, y - x \rangle$. The set of subgradients of $f$ at $x$ is the subderivative or subdifferential of $f$ at $x$ and is denoted $\partial f(x)$.

We shall need the indicator function $\iota_C(x)$, which is zero for $x$ in $C$ and $+\infty$ otherwise, the Fenchel conjugate $f^*(x^*) := \sup_{x} \{\langle x, x^* \rangle - f(x)\}$ and the infimal convolution $f \square g(x) := \inf \{f(y) + g(z) : x = y + z\}$. The central examples of the normal cone to $C$ at $x$ and the distance function $d_C$ are covered by $N_C(x) = \partial \iota_C$ and $d_C = \iota_C \square \| \cdot \|$. We say that a multifunction $T : X \mapsto 2^{X^*}$ is monotone provided that for any $x, y \in X$, $x^* \in T(x)$ and $y^* \in T(y)$,

$$\langle y^* - x^*, y - x \rangle \geq 0,$$

and we say that $T$ is maximal monotone if its graph is not properly included in any other monotone graph. The subdifferential of a convex lower semicontinuous (lsc)
function on a Banach space is a fine example of a maximal monotone multifunction (see [3, 4, 10] wherein other notation and usage may also be followed up).

2. Representative functions

For any monotone mapping $T$, we associate the Fitzpatrick function introduced by Simon Fitzpatrick in [8] but then neglected for many years until re-popularized in papers by Penot [8], Burachik-Svaiter [5], and others. Some more of the related history may be found in [2]. Fitzpatrick’s function is defined on $X \times X^*$ by

$$F_T(x, x^*) := \sup \{ \langle x, y^* \rangle + \langle x^*, y \rangle : y^* \in T(y), y \in \text{dom } T \},$$

which is clearly lower semicontinuous and convex as an affine supremum.

**Proposition 1 ([6, 4]).** For a maximal monotone operator $T$,

$$F_T(x, x^*) \geq \langle x, x^* \rangle$$

with equality if and only if $x^* \in T(x)$.

Correspondingly Penot’s function is also defined on $X \times X^*$, and is given as the convexification

$$P_T(x, x^*) := \inf \left\{ \sum_{i=1}^{N} \lambda_i (x_i, x_i^*) : \sum_{i} \lambda_i (x_i, x_i^*) = (x, x^*), \right. \left. x_i^* \in T(x_i), \sum \lambda_i = 1, \lambda_i \geq 0 \right\}.$$  

It is easy to see that $P_T$ is convex. Moreover, with the conjugate defined on $(X \times X^*)^*$ and canonically restricted to $X \times X^*$ (with the appropriate ordering of variables $x$ and $x^*$), we have

$$P_T^* = F_T \quad \text{while} \quad P_T^{-1} = P_T^* = P_T,$$

where $P_T$ is the lower-semicontinuous hull, the largest lower-semicontinuous minorant, of $P_T$.

More generally, we say that a lower-semicontinuous convex function $H_T$ represents a monotone operator $T$, or is a representative function, if

$$H_T(x, x^*) \geq \langle x, x^* \rangle$$

with equality when $x^* \in T(x)$. We say that a representative is exact if $H_T(x, x^*) = \langle x, x^* \rangle$ exactly on the graph of $T$. Now we may check that:

**Proposition 2 ([6, 8]).** Let $T$ be monotone on a Banach space $X$. Then

1.) Penot’s function $P_T$ represents $T$.

ii.) If $H_T$ represents $T$, then $H_T \leq P_T$ pointwise.

iii.) If $T$ is maximal and $H_T$ represents $T$, then $F_T \leq H_T \leq P_T$.

iv.) $F_T(x, x^*) \leq \langle x, x^* \rangle$ iff $(x, x^*)$ is monotonically related to the graph of $T$.

v.) Suppose $F_T$ represents $T$. Then $F_T(x, x^*) = \langle x, x^* \rangle$ iff $P_T(x, x^*) = \langle x, x^* \rangle$.

**Proof.** (i.) is an easy computation performed in [2, 8].

(ii.) is a direct consequence of $P_T = (c_T)^{**}$ and the fact that $H_T(x, x^*) \leq c_T$, where $c_T(x, x^*) := \langle x, x^* \rangle + \iota_{G_T}(x, x^*)$.

(iii.) The left-hand inequality is established in [6, 8].

(iiv.) is a direct computation.
(v.) By (iv.) and since $F_T$ is representative, we need only show the ‘if’. We observe that if $F_T(x,x') = \langle x,x' \rangle$, then minorizing $F_T(x+t(y-x),x'+t(y'-x'))$ by $(x+t(y-x),x'+t(y'-x'))$ we have

$F_T(y,y') - F_T(x,x') \geq d^+ F_T((y-x),(y'-x')) \geq \langle x,y'-x' \rangle + \langle y-x,x' \rangle$

for all $y,y'$. This shows that $(x',x) \in \partial F_T(x,x')$. Equivalently,

$$2 \langle x,x' \rangle = F_T(x,x') + F_T(x,x') = F_T(x,x') + F_T(x,x')$$

and so $F_T(x,x') = \langle x,x' \rangle$.

Note that $F_T$ need not represent $T$ if $T$ is not maximal. We shall show that the situation is somewhat ameliorated when $T = A + B$ is the sum of maximal monotone operators satisfying

$$0 \in \text{core} \{ \text{conv} D(A) - \text{conv} D(B) \}.$$

To do so, we define two partial infimal convolutions:

$$V_{A,B}(x,x') := \inf \{ F_A(x,u') + F_B(x,v') : u' + v' = x' \},$$

and

$$W_{A,B}(x,x') := \inf \{ P_A(x,u') + P_B(x,v') : u' + v' = x' \}.$$

The first result is very interesting in its own right; it is a lovely observation first exploited by Voisei:

**Theorem 3** (Partial convolution, [12]). Suppose $A$ and $B$ are maximal monotone and satisfy the transversality condition [1]. Then $V_{A,B}(x,x') = W_{A,B}(x,x')$ is norm-weak-star lower-semicontinuous and is attained when these values are finite. Consequently,

$$V_{A,B}(x,x') \geq \langle x,x' \rangle$$

with equality if and only if $x' \in (A + B)(x)$. In particular, $V_{A,B}$ represents $A + B$ and so $V_{A,B} \leq F_{A+B}$.

**Proof.** The argument, based on a conjugate formula of Penot [8 Prop. 13], as in Voiyei [12], and in [2 §5], or a direct Lagrange multiplier calculation [3 §4.3], shows that $V_{A,B}(x,x') = W_{A,B}(x,x')$ and the value is attained when finite. The rest follows since $P_A = F_A$ and $P_B = F_B$ have the representative properties of Proposition [1].

Indeed, $V_{A,B}(x,x') \geq \langle x,x' \rangle$ follows directly from the definition of convolution as does $V_{A,B}(x,x') = \langle x,x' \rangle$ when $x' \in (A + B)(x)$. Finally if $V_{A,B}(x,x') = \langle x,x' \rangle$, we let $x' = u' + v'$ be the attaining values, as assured by the conjugacy formula. Then

$$0 = V_{A,B}(x,x') - \langle x,x' \rangle = \{ F_A(x,u') - \langle x,u' \rangle \} + \{ F_B(x,v') - \langle x,v' \rangle \}.$$

As the bracketed terms are nonnegative we deduce that they are both zero and so $u' \in A(x), v' \in B(x)$ and we are done.

Let us say that a monotone operator $T$ is almost maximal if

$$F_T(x,x') \geq \langle x,x' \rangle$$

for all $x \in X, x' \in X^*$. This is to say that $F_T$ represents $T$. The name is justified since Proposition [1] assures that every maximal monotone operator is almost maximal. Also, if $T$ is maximal and $\text{Gr} \ S = \text{Gr} (T)$ for some other, necessarily monotone, operator $S$, then $S$ is almost maximal.
A nice consequence of the definition is:

**Corollary 4.** Suppose $T$ is almost maximal monotone. Then a closed convex function $\mathcal{H}$ defined on $X \times X^*$ represents $T$ if and only if

$$\mathcal{F}_T \leq \mathcal{H} \leq \mathcal{P}_T.$$  

**Proof.** By Proposition 2 we need only show that each representative function $\mathcal{H}$ is minorized by $\mathcal{F}_T$. Suppose we show that $\mathcal{H}^*$ is also a representative. Then $\mathcal{H}^* \leq \mathcal{P}^*_T \Rightarrow \mathcal{H} \geq \mathcal{F}_T$ as required. To show $\mathcal{H}^*$, restricted to $X \times X^*$, is a representative, since $\mathcal{H}^* \geq \mathcal{F}_T$ we need only show that $\mathcal{H}^*(x,x^*) = \langle x, x^* \rangle$ on $\text{Gr}(T)$. This is the case by the argument of Proposition 2(v.) applied to $\mathcal{H}$. \qed

We offer further justification of the term in the final preliminary result.

**Proposition 5.** Suppose that $A$ and $B$ are maximal monotone operators on a Banach space $X$ and that the transversality condition (1) holds. Then $A + B$ is maximal if and only if it is almost maximal.

**Proof.** Suppose that $(x, x^*)$ is monotonically related to the graph of $A + B$. Then $\mathcal{F}_{A+B}(x, x^*) \leq \langle x, x^* \rangle$. As $A + B$ is almost maximal we deduce that $\mathcal{F}_{A+B}(x, x^*) = \langle x, x^* \rangle$ and so, by part v.) of Proposition 2 we see that $\mathcal{P}_{A+B}(x, x^*) = \langle x, x^* \rangle$. Consequently, an appeal to Theorem 3 shows that $\mathcal{V}_{A+B}(x, x^*) = \langle x, x^* \rangle$ and so that $x^* \in (A + B)(x)$, which completes the proof. \qed

The next corollary shows that, in terms of convergence, $A + B$ shares an important property of maximality. By ‘bdw*’ we denote weak* -convergence for bounded nets (and hence include all weak* -convergent sequences). Thus, the next result is about convergence rather than topology per se.

**Corollary 6** (Graph closedness). Suppose that $A$ and $B$ are maximal monotone in Banach space and that the transversality condition (1) holds. Then $A + B$ has a $\|\cdot\| \times \text{bdw}^*$ closed graph and consequently the images $(A + B)(x)$ are weak* -closed and convex for all $x$ in $X$.

**Proof.** Clearly,

$$\{(x,x^*) : \mathcal{V}_{A,B}(x, x^*) - \langle x, x^* \rangle \leq 0\} = \text{Gr}(A + B)$$

is $\|\cdot\| \times \text{bdw}^*$ closed in the product space, since $\mathcal{V}_{A,B}$ is $\|\cdot\| \times \text{bdw}^*$ lower-semicontinuous while the bilinear form is $\|\cdot\| \times \text{bdw}^*$ continuous. \qed

Dually, we may show that $A + B$ is demi-closed: if $x_n^* \in (A + B)(x_n)$ and $x_n \rightharpoonup x$, while $x_n \rightharpoonup x$ (for sequences or bounded nets), then $x^* \in (A + B)(x)$. Observe that the graph of a maximal monotone operator need not be $\|\cdot\| \times w^*$ closed (or even bw*) even for a subgradient [4 Example 5.2.31]. The $w^*$-closure of $(A + B)(x)$ was first proven in [11], by appealing to the Banach-Dieudonné theorem.

3. **Our main results**

We first provide two useful criteria for almost maximality.

**Proposition 7.** Assume that $S$ is monotone and that either $S$ is surjective or has full domain. Then $S$ is almost maximal.
Assume also that \( D \) on a Banach space. Suppose also that either \( A \) is maximal monotone. Then
\[
\mathcal{F}_S(x, x^*) \geq \langle x, s^* \rangle + \langle s, x^* \rangle - \langle s, s^* \rangle = \langle x, x^* \rangle.
\]
The other case is similar. \( \square \)

We denote the \textit{algebraic closure} of a set at \( x \in C \) by
\[
C^{alg}(x) := \{ d : \exists t_n < 1, t_n \to 1, \text{ such that } t_n d + (1 - t_n)x \in C \}.
\]
We write \( C^{alg} := \bigcap_{x \in C} C^{alg}(x) \).

**Proposition 8.** Assume that \( A \) and \( B \) are maximal monotone and that \( A \) holds. Assume also that \( D(A) \cap D(B) \) is closed and
\[
(2) \quad \text{conv} \ D(A) \cap \text{conv} \ D(B) = \overline{D(A) \cap D(B)}^{alg}.
\]
Then \( A + B \) is almost maximal.

**Proof.** Assume that \( 0 \in D(A) \cap D(B) \). Let \( U := \text{conv} \ D(A) \) and \( V := \text{conv} \ D(B) \). Note also that \( A \) implies that the normal cone satisfies \( N_U + N_V = N_{U \cup V} \). Much as in \( [12] \) we argue by maximality that \( A = A + N_U \) and \( B = B + N_V \). Thus, using \( [2] \) shows that
\[
A + B = (A + B) + N_{D(A+B)}.
\]

Now suppose that \( \mathcal{F}_{A+B}(x, x^*) \leq \langle x, x^* \rangle \). This implies that for every \( n^* \in N_{D(A+B)}(y) \) and \( s^* \in (A + B)(y) \) we have
\[
\langle s^* + tn^* - x^*, y - x \rangle \geq 0,
\]
for all \( t > 0 \). Hence \( 0 \in N_{D(A+B)}(x) \) by maximality of the normal cone since \( D(A + B) \) is closed.

Thus, \( x \in \overline{D(A+B)} \). Now \( [2] \) implies that \( x \) lies in \( \overline{D(A+B)}^{alg} \). Thus we can select \( z^* \in (A + B)(\alpha x) \), for \( 0 < \alpha < 1 \). Then, by convexity and by definition,
\[
\alpha \mathcal{F}_{A+B}(x, x^*) = \alpha \mathcal{F}_{A+B}(x, x^*) + (1 - \alpha)\mathcal{F}_{A+B}(0, 0) \geq \mathcal{F}_{A+B}(\alpha x, \alpha x^*) \geq \langle \alpha x, \alpha x^* \rangle + \langle \alpha x, z^* \rangle - \langle \alpha x, z^* \rangle = \alpha^2 \langle x, x^* \rangle.
\]

Thus \( \mathcal{F}_{A+B}(x, x^*) \geq \alpha \langle x, x^* \rangle \). Letting \( \alpha \uparrow 1 \) completes the proof. \( \square \)

**Theorem 9** (Maximality of sums, I). Suppose that \( A \) and \( B \) are maximal monotone on a Banach space. Suppose also that either

i.) the set \( \text{int} \ D(A) \cap \text{int} \ D(B) \) is nonempty; or

ii.) \( D(A) \cap \text{int} \ D(B) \neq \emptyset \) while \( D(A) \cap D(B) \) is closed and convex; or

iii.) both \( D(A) \), \( D(B) \) are closed and convex and \( 0 \in \text{core conv} \ \{ D(A) - D(B) \} \).

Then \( A + B \) is maximal monotone.

**Proof.** Each of the hypotheses leads to \( \text{conv} \ D(A) \cap \text{conv} \ D(B) \subset \overline{D(A+B)}^{alg} \), since \( D(A) \) is convex when \( D(A) \) has nonempty interior; see \( [2] \ [9] \ [10] \). Moreover the hypotheses of i.) and ii.) imply \( [1] \). Thus, Proposition 8 applies as then does Proposition 5 as soon as \( D(A + B) \) is closed. \( \square \)
Part iii.) of Theorem [9] is the main result in [12]. In [13] corresponding results are given for compositions with closed convex domain. Our results herein have a corresponding extension. A more general proof of Theorem [9] i.) follows from results in [2].

**Theorem 10** (Maximality of sums, II). Suppose A and B are maximal monotone on a Banach space. Suppose also that \( \text{core conv}(A) \cap \text{core conv}(B) \neq \emptyset \). Then \( A + B \) is maximal monotone.

**Proof.** Suppose \((x, x^*)\) is monotonically related to the graph of \( A + B \). Let \( W \) be an arbitrary basic weak-star zero neighbourhood. Fix a finite-dimensional subspace \( F \) of \( X \) containing both \( (x, x^*) \) and \( A + B \) we observe that \((x, x^*)_F\) is monotonically related to the graph of \((A + B)_F\). Hence by maximality \( x^*_F \in (A + B)_F(x) \). Consequently,

\[
x^* \in (A + B)(x) + F^\perp \subset (A + B)(x) + W.
\]

Since \( W \) is arbitrary, applying Corollary [6] we deduce that

\[
x^* \in (A + B)(x)^* = (A + B)(x),
\]

by the Veronas’ part of Corollary [6] and we are done. \( \square \)

**Remark 11.** This argument works under the Brezis-Attouch condition [11] if we can ensure that for each finite-dimensional space \( F \) there is a reflexive superspace, \( R \), such that \( A_R \) and \( B_R \) are both maximal. This is the case, for example, if after translation, \( 0 \in D(A) \cap \text{core conv}(B) \neq \emptyset \) and for each finite-dimensional \( F \) there is a reflexive subspace \( R \) containing \( F \) such that \( A|_R \) is maximal. Thus, any counterexample to the sum theorem has to have quite messy domains. Another proof of this result can probably be obtained from Asplund’s decomposition of a maximal monotone operator as the sum of cyclic and acyclic operators, as described in [2] §3. Note also that maximality of \( T_Y \) and that of \( T + \partial_f \) are equivalent for a closed subspace \( Y \).

**Remark 12.** Suppose \( f \) is lower-semicontinuous, proper and convex in a Banach space. We observe, without appealing to maximality, that the representative function \((f \oplus f^*)(x, x^*) := f(x) + f^*(x^*)\) coincides with \( \langle x, x^* \rangle \) exactly for \( x^* \) in \( \partial f(x) \). As in Proposition [5] to prove that \( \partial f \) is maximal it thus suffices to show that \( \partial f \) is almost maximal. Can this be done any more efficiently than directly proving maximality via an approximate mean-value theorem, as say in [4] Thm. 3.4.6? \( \square \)

**Remark 13.** We can probably significantly improve the result in the case where one operator is a subgradient because the representative function \((f \oplus f^*)(x, x^*) := f(x) + f^*(x^*)\) is exact. We define

\[
\mathcal{F}_{T,f}(x, x^*) := f(x) + \sup_{y^* \in T(y)} \{ \langle x, y^* \rangle + \langle y, x^* \rangle - \langle y, y^* \rangle - f(y) \},
\]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
with conjugate
\[ P_{T,f}(x, x^*) := f(x) + \text{conv}_{y^i \in T(y_i)} \{ y_i^* - f(y_i) \} \]
and note that \( F_{T,0} = F_T \) and \( P_{T,0} = \overline{P_T} \), and, as before, \( P_{T,f} = F_{T,f} \leq P_{T,f} \).

We conjecture that \( \langle x, x^* \rangle \) monotonically related to \( \text{Gr}(T + \partial f) \) as holds in the extreme cases where one of \( T \) or \( f \) is zero.

**Theorem 14** (Maximality of sums, III). Suppose \( T \) is maximal monotone and \( f \) is convex and closed. Suppose that (3) holds and that
\[ \text{dom } f \cap \text{core conv } D(T) \]
is nonempty. Then \( T + \partial f \) is maximal, whenever it has closed domain.

**Proof.** We define
\[ \mathcal{V}_{T,f}(x, x^*) := f(x) + (F_{T,f}(x, \cdot)\square f^*)(x^*) \]
Then
\[ \mathcal{V}_{T,f}(x, x^*) \leq P_{T,f}(x, x^*) \]
By assumption \( F_{T,f}(x, x^*) \leq \langle x, x^* \rangle \) for \( (x, x^*) \) monotonically related to \( \text{Gr}(T + \partial f) \) while, as in Proposition 3, \( \mathcal{V}_{T,f}(x, x^*) \) is an exact representative for \( T + \partial f \). Now the weakened constraint qualification (4) still ensures that \( F_{T,f} \) represents \( T + \partial f \) as it implies that \( \text{dom } f \cap \text{conv } D(T) \subseteq \text{dom } f \cap \text{conv } D(T)^{\text{alg}} \). Much as before,
\[ F_{T,f}(x, x^*) = \langle x, x^* \rangle \Rightarrow P_{T,f}(x, x^*) = \langle x, x^* \rangle \Rightarrow \mathcal{V}_{T,f}(x, x^*) = \langle x, x^* \rangle \]
and we are done. \( \square \)

We finish with two especially nice consequences of part ii.) of Theorem 9.

**Corollary 15** (Normal cones). Suppose in an arbitrary Banach space that \( T \) is maximal monotone and \( C \) is closed and convex while \( C \cap \text{int } D(T) \neq \emptyset \). Then \( T + N_C \) is maximal monotone, if either \( D(T) \) is closed or \( C \) is non-empty.

**Proof.** The maximality of \( T + N_C \) is an immediate consequence of part ii.) of Theorem 9 or of Theorem 10. \( \square \)

Recall that a maximal monotone mapping \( T \) is maximal monotone locally [11], or of type (FPV), if for every open convex set \( V \) in \( X \) with \( V \cap D(T) \neq \emptyset \) the following holds for every \( x \in V : \langle y^*-x^*, y-x \rangle \geq 0 \) for all \( y^* \in T(y) \) and all \( y \in V \) implies that \( x^* \in T(x) \).

**Corollary 16.** Suppose in an arbitrary Banach space that \( T \) is maximal monotone and \( D(T) \) has nonempty interior, or is closed and convex. Then \( T \) is of type (FPV).

**Proof.** We argue as follows. Fix \( x, V \) and \( x^* \) as in the definition of (FPV). We may select a closed convex set \( C \) such that \( x \in \text{int } C \subset V \) and \( \text{int } D(T) \cap \text{int } C \neq \emptyset \). It follows from Corollary 9 that \( T+N_C \) is maximal. Let \( y^* \in T(y), n^* \in N_C(y), y \in Y \) be given. Then \( \langle y^* + n^*-x^*, y-x \rangle = \langle y^*-x^*, y-x \rangle + \langle n^*, y-x \rangle \geq 0 \) since \( x \in C \). By maximality \( x^* \in T(x) + N_C(x) = T(x) \) since \( x \in \text{int } C \).

The same argument shows that every maximal monotone mapping with closed convex domain is type (FPV). \( \square \)
The case $D(T) = X$ of Corollary 16 was first established in [7].

4. Final remarks

As distinct from the reflexive case, we note that our arguments make no use of the duality map. Indeed, outside of reflexive space, the Rockafellar-Minty approach via surjectivity of $T + J$ is of little use; see [2, §6] and [1, §5.1]. That said, Theorem 9 certainly suggests that $A + B$ may well be maximal given only (1) and no auxiliary conditions.

References


Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada
E-mail address: jborwein@cs.dal.ca