A General Sum Theorem for Maximality of Monotone Operators

A. C. Eberhard and J. M. Borwein*

October 28, 2007

Abstract

We establish maximality of the sum of two maximal monotone operators in general Banach space, assuming only the Rockafellar qualification assumption. This resolves, affirmatively, a thirty-five year old conjecture.

1 Introduction

The theory of maximal monotone operators in reflexive space is rather well understood and complete, as illustrated by Rockafellar’s foundational work in [19]. A good indication of our understanding fifteen years ago can be gathered from Phelps’ expositions in [17] and [18]. Outside of reflexive space our knowledge is more fragmentary. This reflects on the difficulty of the subject (see its early roots in [11, 12]) and more especially on the paucity of tools available until recently—in reflexive space key use was made of the surjectivity of the duality map. The situation was ameliorated significantly by Simons’ monograph [23] which made more central the role of the Hahn-Banach theorem, and by the rediscovery of Fitzpatrick’s 1988 paper [10]—beautifully exploited in Penot’s work [15] and that of Simons [23] and others.

As described in more detail in [3], this has allowed for a reduction of much monotone operator theory to convex analysis, culminating in very clean and elementary proofs of the maximality of the sum in reflexive space [3, 7, 26] under the weakest-known constraint qualification. It has also allowed for a flowering of results in non-reflexive space, most recently in conditions for the sum to be maximal using an approach pioneered by Voisei [28, 29, 5].

In this paper we build upon this foundation to supply various results establishing, in Theorem 31, that in any Banach space the sum of two maximal

*The second author’s research was funded by NSERC and the Canada Research Chair programme, and the first author’s by ARC grant number DP0664423. This study was commenced between August and December 2005 while the first author was visiting Dalhousie University.

2000 Mathematics Subject Classification: 47H05, 46N10, 47H04, 46A20, 49J53.

Keywords and Phrases: Maximal monotone operators, convex analysis, Fitzpatrick function, Fenchel duality, sum theorem.
monotone operators $T$ and $M$ is maximal under Rockafellar’s original condition:

$$\emptyset \neq \text{dom } M \cap \text{int } (\text{dom } T).$$

A particularly pleasant consequence is given in Corollary 34: the domain of every maximal monotone operator is semi-convex, i.e., with convex closure.

## 2 Notations and Preliminaries

Suppose $Z$ is a Banach space and $Z^*$ is its dual. Denote the open ball in the primal space by $B_r(z) = \{ z \mid \| z - z_0 \| \leq r \}$ and similarly in the dual space denote the open ball by $B^*_r(z^*) = \{ z^* \mid \| z^* - z^* \| \leq r \}$. We may view $Z \times Z^*$ paired with $Z^* \times Z$ using the coupling $\langle (z, z^*), (x, x^*) \rangle = \langle z, x^* \rangle + \langle z, x^* \rangle$ and the norm $\|(z, z^*)\|^2 = \|z\|^2 + \|z^*\|^2$. At times we will assume $Z \times Z^*$ is endowed with the product topology $s \times \text{bw}^* (Z, Z^*)$ formed from the strong topology on $Z$ and the bounded weak* on $Z^*$ (see [14] page 150–154). We will pair this space with $Z^* \times Z$ which is endowed with product topology $\text{bw}^* \times s (Z^*, Z)$ when $Z^*$ is endowed with the bounded weak* topology and $Z$ the strong topology. This is a valid pairing due to the fact that when $Z$ is Banach the $\text{bw}^*$-continuous linear functionals on $Z^*$ are canonically isomorphic to $Z$. When $Z = X^*$ we obtain the pairing of $s \times \text{bw}^* (X^*, X^{**})$ with $\text{bw}^* \times s (X^{**} \times X^*)$, which is of interest in the study of maximal monotone operators using representative functions. When we pair the space $s \times \text{bw}^* (Z, Z^*)$ with $\text{bw}^* \times s (Z^*, Z)$ the associated convex conjugation operation of a proper convex function $f \in \Gamma_{s \times \text{bw}^*} (Z, Z^*)$ is denoted by $f^* \in \Gamma_{\text{bw}^* \times s} (Z^*, Z)$. When we associate $X \times X^*$ with $X^* \times X^{**}$ using the weak topology on $X \times X^*$ and the weak* on $X^* \times X^{**}$ we denote the conjugate of $f \in \Gamma_w (X, X^*)$ to be $\hat{f}^* \in \Gamma_{w^*} (X^*, X^{**})$.

When using the pairing alluded to above then we may consider $f^* (x^*, x) := \hat{f}^* (x^*, J_X (x))$ (tolerating the abuse of notation). The *epigraph* of $f$ is the set $\text{epi } f := \{(x, x^*, \alpha) \in X \times X^* \times \mathbb{R} \mid \alpha \geq f(x, x^*)\}$. The *indicator function* of a set $T \subseteq X \times X^*$ is denoted by $\delta_T$. The *support function* of a set $A \subseteq X \times X^*$ is given by $\delta_A^* (x^*, x)$. When we pair the spaces $X \times X^*$ with $X^* \times X$, the second conjugate $f^{**} (x, x^*) = f (x, x^*)$ whenever $f$ is a jointly strong-bounded weak* continuous, proper convex function. Occasionally we will use the shorthand notation $J_X (X) = \hat{X} \subseteq X^{**}$ along with $J_X (x) = \hat{x}$ to denote the embedding into $X^{**}$. Alternatively, one can view $f^{**} (x, x^*) = \left( \hat{f}^* \right)^* (J_X (x), x^*)$.

To simplify notation for a multi-function $T : X \rightrightarrows Y$ we denote its graph by $T := \{(x, y) \in X \times Y \mid y \in T(x)\}$. We say $T$ is a monotone set if

$$\forall (x, x^*) \in T \ \forall (y, y^*) \in T \ \langle x - y, x^* - y^* \rangle \geq 0.$$

(2)

If $T$ does not possess a proper monotone extension then $T$ is said to be maximal monotone. We say $(y, y^*)$ is monotonically related to $T$ when $(\forall (x, x^*) \in T)$ we have $\langle x - y, x^* - y^* \rangle \geq 0$. When $T$ is maximal then $(y, y^*) \notin T$ implies the existence of $(x, x^*) \in T$ such that $\langle x - y, x^* - y^* \rangle < 0$. 2
2.1 Representative Functions

**Definition 1** The Fitzpatrick function associated with an operator $T : X \rightrightarrows X^*$ is defined by

$$\mathcal{F}_T (y, y^*) := \sup_{(x, x^*) \in T} \{ \langle y, x^* \rangle + \langle x, y^* \rangle - \langle x, x^* \rangle \}$$

$$= \left\{ \langle y, y^* \rangle - \inf_{(x, x^*) \in T} \langle y - x, y^* - x^* \rangle \right\}$$  \hspace{1cm} (3)

As $(y, y^*) \mapsto \langle y, x^* \rangle + \langle x, y^* \rangle - \langle x, x^* \rangle$ is continuous with respect to the strong-weak* product topology we find that $\mathcal{F}_T (y, y^*)$ is a jointly strong-weak* lower semi–continuous, proper convex function whenever $T$ is monotone. Alternatively one may observe that $(y, y^*) \mapsto \langle y, x^* \rangle + \langle x, y^* \rangle - \langle x, x^* \rangle$ is continuous with respect to the strong topology on $X \times X^*$ and state that $\mathcal{F}_T (y, y^*)$ is a jointly strongly lower semi–continuous proper convex function on $X \times X^*$. As the strong and weak closures of a convex body coincide then we also may state that $\text{epi} \mathcal{F}_T$ is a weakly closed, convex subset of $X \times X^*$. From definition (3) it is easily seen that for $y^* \in T(y)$ we have

$$\mathcal{F}_T (y, y^*) = \langle y, y^* \rangle < +\infty,$$  \hspace{1cm} (4)

when $T$ is monotone. Next, $\mathcal{F}_T (y, y^*) \geq \langle y, y^* \rangle$ holds for all $(y, y^*)$ with equality when $y^* \in T(y)$ if and only if $T$ is maximal monotone, i.e.,

$$T = \{ (y, y^*) \mid \mathcal{F}_T (y, y^*) = \langle y, y^* \rangle \}.$$

Define the ‘transpose’ operator $\dagger$ : $(x^*, x) \mapsto (x, x^*)$ and $c_T (\cdot, \cdot) := \delta_T (\cdot, \cdot) + \langle \cdot, \cdot \rangle$, with conjugate

$$c_T^* (y, y^*) := \sup_{(w, w^*) \in X \times X^*} \{ \langle y, w^* \rangle + \langle w, y^* \rangle - \langle \delta_T (w, w^*) + \langle w, w^* \rangle \}$$

$$= \sup_{(x, x^*) \in T} \{ \langle y, w^* \rangle + \langle w, y^* \rangle - \langle w, w^* \rangle \} = \mathcal{F}_A (y, y^*).$$

The second conjugate of $c_T$ is of interest in that it leads to the Penot representative function:

$$\mathcal{P}_T (y, y^*) := \mathcal{F}_T^* (y, y^*) := (c_T^*)^* (y, y^*) = \mathcal{P}_T (y, y^*)$$

where

$$\mathcal{P}_T (y, y^*) = \inf \left\{ \sum_i \lambda_i (x_i, x_i^*) \mid \sum_i \lambda_i (x_i, x_i^*, 1) = (y, y^*, 1), \right\}$$

$$(x_i, x_i^*) \in T, \lambda_i \geq 0 \}.$$  \hspace{1cm} (5)

where the closure may be taken with respect to the topology on $X \times X^*$ consistent with the pairing. Clearly we have $\mathcal{P}_T (y, y^*) = +\infty$ if $(y, y^*) \notin \text{co} T$ and $\text{co} T \subseteq \text{dom} \mathcal{P}_T$.

We recall that if $\text{core} \text{dom} T \neq \emptyset$ and $T$ is maximal then both $\text{dom} \overline{T}$ and $\text{int} \text{dom} T$ are convex [3]. In a reflexive space maximality is sufficient to ensure $\text{dom} T$ is convex, a property referred to as semi–convexity.
Recall that a representative function of a monotone mapping $T$ on $X$ is a convex function $\mathcal{H}_T$ on $X \times X^*$ such that $\mathcal{H}_T(y, y^*) \geq \langle y, y^* \rangle$ for all $(y, y^*) \in X \times X^*$ with $\mathcal{H}_T(y, y^*) = \langle y, y^* \rangle$ when $y^* \in T(y)$. When $T$ is not specified we say a $s$-bw* $(X, X^*)$ closed, proper convex function $f$ is representative when $f(y, y^*) \geq \langle y, y^* \rangle$ for all $(y, y^*) \in X \times X^*$. It has been noted by a number of authors, see \cite{5}, that when $T$ is maximal the largest representative function is $\mathcal{P}_T$ and the smallest is $\mathcal{F}_T$ (and so $\mathcal{P}_T \geq \mathcal{F}_T$) pointwise. Denote the restriction of a subset $U \subseteq X^* \times X^*$ to the space $X^* \times X$ by

$$R_{X^* \times X}(U) := U \cap (X^* \times X).$$

For a subset $V \subseteq X^* \times X$ denoted by $J_{X^* \times X}(V)$ the imbedding of $V$ into $X^* \times X^*$. Occasionally we will use the shorthand notation $J_X(X) = \tilde{X}$ along with $J_X(x) = \tilde{x}$ to denote the embedding into $X^*$. We may pair the space $s \times$ bw* $(Z, Z^*)$ with bw* $s (Z^*, Z)$ and when this is done we denote the associated convex conjugation operation of a proper convex function $f \in \Gamma_{s \times \text{bw}*} (Z, Z^*)$ by $f^* \in \Gamma_{\text{bw}* \times s} (Z^*, Z)$.

Fitzpatrick \cite{10} showed that

$$\text{epi} \mathcal{P}_T = \left[ R_{X^* \times X \times \mathbb{R}} \left( \text{epi} \hat{\mathcal{F}}_T \right) \right]^\dagger \quad \text{and epi} \mathcal{F}_T = \left[ R_{X^* \times X \times \mathbb{R}} \left( \text{epi} \hat{\mathcal{P}}_T \right) \right]^\dagger$$

are representative functions when $T$ is maximal monotone, and in \cite{3} it is shown that $\mathcal{P}_T$ is representative when $T$ is just monotone.

**Lemma 2** (\cite{3}) *For any monotone mapping $T$ the function $\mathcal{P}_T : X \times X^* \to \mathbb{R}$ is a representative convex function for $T$.***

Because $f \leq \langle \cdot, \cdot \rangle + \delta_T = c_M$ we have $c_T^* (x^*, x^*) \leq f^*$. Thus, when $T$ is maximal

$$\langle x, x^* \rangle \leq \mathcal{F}_T(x, x^*) \leq \hat{f}^* (x^*, \tilde{x}) = f^* (x^*, x).$$

That is, when $f$ is a representative function then the embedding of $\hat{f}^* (x^*, x^*)$ into $X^* \times X$ is also a representative function. Hence

$$T : X \to X^* \quad \begin{cases} f \text{ representative} \end{cases} \quad \Rightarrow \quad f(x, x^*) \geq \langle x, x^* \rangle, \quad \forall (x, x^*) \in X \times X^* \quad \begin{cases} f^*(x^*, x) \geq \langle x^*, x \rangle, \quad \forall (x, x^*) \in X \times X^* \end{cases}$$

(6)

It is natural to ask if the full conjugate $\hat{f}^* (x^*, x^*)$ is also representative on $X^* \times X^*$. This question has only recently been answered in \cite{9}.

**Theorem 3** (\cite{9}) *Suppose $T$ is a maximal monotone operator on $X \times X^*$. Then $\hat{\mathcal{P}}_T^*$ and $\hat{\mathcal{F}}_T^*$ are representative functions on $X^* \times X^*$.***

In \cite[Prop. 1]{16} and \cite{2} the converse to this result was proved in a reflexive space. That is, when the inequalities on the right of (6) hold then $T := \{ (x, x^*) \mid f(x, x^*) = \langle x, x^* \rangle \}$ is a maximal monotone set. This provides
a complete and simple characterisation of maximality in the context of reflexive spaces. This fact will be used a number of times in this paper. In passing we note that when $F_T$ is a representative function then—as $P_T$ is always representative—the above considerations apply. One should note that in this case $M := \{(x, x^*) | F_T(x, x^*) = \langle x, x^* \rangle \}$ is maximal when $X$ is reflexive but we may have $M \neq T$ (consider the case when only the closure of $T$ equals $M$).

When $F_T$ is a representative function we say that $T$ is almost maximal.

Another construction that arises in this paper occurs when we embed the graph of a monotone mapping $T \subseteq X \times X^*$ into $X^{**} \times X^*$. Then it is possible to obtain a maximal monotone extension $\tilde{T}$ of the embedded $J_{X \times X^*}(T) = \tilde{T} \subseteq X^{**} \times X^*$. An interesting question arises as to when the restriction $R_{X \times X^*}(\tilde{T})$ coincides with $T$. When dealing with a maximal monotone operator $\tilde{T} \subseteq X^{**} \times X^*$ it is possible to consider both $P_{\tilde{T}}$ and $F_{\tilde{T}} : X^{**} \times X^* \to \mathbb{R}$ as a representative functions which would then be consider to be bw*-s $(X^{**}, X^*)$ closed, proper convex functions.

2.2 Structure of the Paper and Flow of Logic

We now outline the flow of logic that leads to our main theorem regarding sums. The reader may find it useful to consult this outline as a road-map, since several technical results are required before a proof of the main sum theorem is possible.

In section 3 we make some critical observations regarding the maximality of certain extensions of a monotone set $T$ when it is embedded into the dual space $X^* \times X^{**}$. These results can be viewed as a generalization of the results [16, Prop. 1] to non-reflexive spaces. This allows us to prove Theorem 8 which provides sufficient conditions (formally weaker than maximality) to ensure that the Fitzpatrick function $F_T$ is a representative function.

The strategy leading to the sum Theorem 31, and consequences, is to prove that the conditions of Theorem 8 are satisfied by $T + M$ which will allow us to deduce $F_{T+M}$ is a representative function. The proof is then completed as in [5] by showing that $T + M$ is maximal. In Sections 5 we work as follows:

1. To apply Theorem 8 we need to identify a representative function for $T + M$ to which the Theorem applies. Thus, we choose $h$ as in Theorem 22 and hence require (29) to hold for the choices of $f := F_T$ and $g := F_M$.

2. The second condition requiring validation is: $h^*(x^*, x^{**}) \geq \langle x^{**}, x^* \rangle$ for all $(x^*, x^{**}) \in X^* \times X^{**}$. This follows from Theorem 21, along with the equalities $f^* = F_T^* \geq \langle \cdot, \cdot \rangle$ and $g^* = F_M^* \geq \langle \cdot, \cdot \rangle$ valid on $X^* \times X^{**}$ (as observed in Theorem 3) assuming the qualification (28).

3. The final condition needed in Theorem 8 is that there exists a family of finite-dimensional subspaces that cover $X$ such that $(T + M)_A := A^* \circ (T + M) \circ A$ is maximal for $A : Y \to X$ (the embedding of $Y$ in $X$).

4. As $Y$ is finite-dimensional, showing maximality of $(T + M)_A$ reduces to prescription of a representative function, we denote by $H$, that has the
properties $H \geq \langle \cdot, \cdot \rangle$ and $H^* \geq \langle \cdot, \cdot \rangle$ on $Y \times Y^*$ (see [16, Prop. 1]). Consequently we define $H$ as in Proposition 23 and note that we have the intermediate conjugation formula as described in Proposition 23.

The inequality $H \geq \langle \cdot, \cdot \rangle$ is immediate from Theorem 21 and the fact that $h$ is a representative function for $(T + M)$, by Theorem 22. Thus, all that remains to show $H$ is a representative function of a maximal monotone operator is to show that $H^* \geq \langle \cdot, \cdot \rangle$ on $Y \times Y^*$. The idea is then to only impose an interiority condition on $\text{dom } T$ and leave $\text{dom } M$ general. Then we apply the duality given in Proposition 23, from which we can argue as in Proposition 24 to establish the inequality $H^* \geq \langle \cdot, \cdot \rangle$ on $Y \times Y^*$ (which requires Theorem 10). This establishes that $H$ is indeed a representative function of a maximal monotone operator on $Y$, which we denote by $Q_A$.

5. Finally, we are finished once we can identify $H$ as a representative function of $(T + M)_A$. The trouble is that now one cannot assume that the infimum in the definition of $H$ is achieved. The best one can do is to argue as in Proposition 25 and so to obtain an inclusion inside $\text{int dom } T \cap \hat{Y}$.

6. This inclusion needs to be extended to the $\text{dom } Q_A$ where $Q_A$ is the maximal monotone set defined by $H$. Here Corollary 19 comes to the rescue, allowing us to take limits and add normal cones to reconstruct $Q_A$ from information within $\text{ri dom } Q_A$. This requires us to make sure the normal cones match and that the domain closures of $Q_A$ and $(T + M)_A$ are the same. This is shown in a number of Lemmas in section 4.

All this is required because the obvious shortcut is to impose an interiority condition on both $\text{dom } T$ and $\text{dom } M$, as in [5], which would not cover Rockafellar’s qualification.

3 Conditions for Maximality in a Banach Space

In this section we derive conditions that may be used to deduce maximality of monotone sets defined by certain representative functions. The proof of the sum theorem is based on these results. Thus, we present the technical machinery in this section with the consequences obtained in later sections. Proofs of some recent results are included for the readers convenience. The following is a consequence of Rockafellar’s version of the duality theorem.

**Theorem 4 ([22])** Let $X$ be a normed space, $f : X \to \mathbb{R}$ be a proper and convex function, $g : X \to \mathbb{R}$ be convex and continuous and suppose $f + g \geq \alpha$ on $X$. Then there exists $x^* \in X^*$ such that $f^* (x^*) + g (-x^*) \leq -\alpha$.

For $(x, x^*) \in X \times X^*$ denote
\[
\Delta(x, x^*) = \frac{1}{2} \|x\|^2 + \langle x, x^* \rangle + \frac{1}{2} \|x^*\|^2 \\
\geq \frac{1}{2} \left( \|x\|^2 - 2 \|x\| \|x^*\| + \|x^*\|^2 \right) \geq (\|x\| - \|x^*\|)^2 \geq 0.
\]
Similarly for \((x^{**}, x^*) \in X \times X^*\) denote \(\Delta(x^{**}, x^*) = \frac{1}{2} \|x^{**}\|^2 + \langle x^{**}, x^* \rangle + \frac{1}{2} \|x^*\|^2\).

On \(X \times X^*\) we use the norm \(\| (x, x^*) \|^2 := \|x\|^2 + \|x^*\|^2\) and on \(X \times X^*\) we use \(\|(x^{**}, x^*)\|^2 := \|x^{**}\|^2 + \|x^*\|^2\), since then

\[
\left(\frac{1}{2} \| \cdot \|^2\right)(x^{**}, x^*) = \frac{1}{2} \left(\|x^*\|^2 + \|x^{**}\|^2\right) = \frac{1}{2} \|(x^{**}, x^*)\|^2.
\]

(7)

**Lemma 5** ([26], Lemma 1.3) Let \(h : X \times X^* \to \mathbb{R}\) be proper, convex and \((w, w^*) \in X \times X^*\) and

\[(x, x^*) \in X \times X^* \Rightarrow h(x, x^*) - \langle x, x^* \rangle + \Delta(w - x, w^* - x^*) \geq 0.\]

(8)

Then there exists \((x^*, x^{**}) \in X^* \times X^{**}\) such that

\[
\hat{h}^*(x^*, x^{**}) - \langle x, x^* \rangle + \Delta(w - x^{**}, w^* - x^*) \leq 0.
\]

(9)

**Proof.** Let \(\eta_{(w, w^*)}^{(x, x^*)} = -\langle x, x^* \rangle + \Delta(w - x, w^* - x^*)\). Then since

\[
\eta_{(w, w^*)}^{(x, x^*)}(w, w^*) - \langle x, x^* \rangle = \frac{1}{2} \|(w, w^*) - (x, x^*)\|^2
\]

we have \((x, x^*) \mapsto \eta_{(w, w^*)}^{(x, x^*)}\) is convex and norm-continuous. Then

\[
\eta^*_{(w, w^*)}(x^* - x^{**}) = \sup_{(z, z^*) \in X \times X^*} \left\{ \langle (z, z^*), (x^{**}, -x^*) \rangle - \left( \langle w, w^* \rangle - \langle (z, z^*), (w, w^*) \rangle + \frac{1}{2} \|(w, w^*) - (z, z^*)\|^2 \right) \right\}
\]

\[
= \langle w, w^* \rangle - \langle (x^{**}, x^*), (w, w^*) \rangle + \sup_{(z, z^*) \in X \times X^*} \left\{ \langle (z, z^*) - (w, w^*), (w, w^*) - (x^{**}, x^*) \rangle - \frac{1}{2} \|(w, w^*) - (z, z^*)\|^2 \right\}
\]

\[
= \langle w, w^* \rangle - \langle (x^{**}, x^*), (w, w^*) \rangle + \frac{1}{2} \|(w, w^*) - (x^{**}, x^*)\|^2 = \eta_{(w, w^*)}^{(x, x^*)}(x^{**}, x^*).
\]

Note that this only relies on the identity (7). Now invoke Theorem 4 with the property that \(f(x, x^*) \geq (x, x^*)\) for all \((x, x^*) \in X \times X^*\). As (8) implies \(\inf_{X \times X^*} \{ \hat{h}(x, x^*) + \eta_{(w, w^*)}^{(x, x^*)}\} \geq 0\) we deduce that

\[
\min_{X \times X^{**}} \left\{ \hat{h}^*(x^*, x^{**}) + \eta^*_{(w, w^*)}(x^{**}, x^*) \right\} \leq 0.
\]

This clearly implies (9).

When \(M \subseteq X \times X^*\) is a maximal monotone set we can always embed

\[
J_{X \times X^*}(M) \subseteq X^{**} \times X^*
\]
and extend \( J_{X \times X^*}(M) \) to a larger monotone set \( \widetilde{M} \) within \( X^{**} \times X^* \) (many such extensions have been proposed in the literature).

This set has the following useful property: When \((\tilde{u}, u^*)\) is monotonically related to \( \widetilde{M} \), as it must then be monotonically related to \( M \) as well, we have \((\tilde{u}, u^*)\) \( \in \widetilde{M} \). We will see that this property is more widely held by monotone sets \( M \) which we do not know \textit{a priori} are maximal. Note also that when we choose \( M \) to be a maximal extension of \( J_{X}(M) \) within \( X^{**} \times X^* \) then its Fitzpatrick function satisfies \( \mathcal{F}_M(x^*, x^*) \geq \langle x^*, x^* \rangle \) for all \((x^*, x^*) \in X^{**} \times X^* \) and consequently also \( \mathcal{F}_{\widetilde{M}}(\tilde{x}, x^*) \geq \langle \tilde{x}, x^* \rangle \). By [15, Prop. 4] we know that the following is a monotone set:

\[
\{ (x, x^*) \mid \mathcal{F}_M(x, x^*) = \langle \tilde{x}, x^* \rangle \}.
\]

Clearly \( \mathcal{F}_M(\tilde{x}, x^*) \leq \langle \tilde{x}, x^* \rangle \) if and only of \((\tilde{x}, x^*)\) is monotonically related to \( \widetilde{M} \) and as mentioned earlier it follows that \((x, x^*) \in M \). Thus,

\[
\{ (x, x^*) \mid \mathcal{F}_M(x, x^*) = \langle \tilde{x}, x^* \rangle \} \subseteq M.
\]

Note that as \( \widetilde{M} \) is maximal

\[
\{ (x^*, x^*) \mid \mathcal{F}_M(x^*, x^*) = \langle x^*, x^* \rangle \} = \widetilde{M} \supseteq J_{X \times X^*}(M)
\]

and so

\[
\{ (x, x^*) \mid \mathcal{F}_M(x, x^*) = \langle \tilde{x}, x^* \rangle \} = \widetilde{M} \cap (J_{X}(X) \times X^*) \supseteq J_{X}(M).
\]

Thus \( \mathcal{F}_{\widetilde{M}} \) restricted to \( X \times X^* \) is a representative function of \( M \) and \( M \) is the restriction to \( X \times X^* \) of a maximal monotone set \( \widetilde{M} \) in \( X^{**} \times X^* \). We will now study this phenomenon in more detail.

Denote

\[
M_f := \{(x, x^*) \in X \times X^* \mid f(x, x^*) = \langle x, x^* \rangle \}
\]

\[
M_{f^*} := \{x^* \in X^* \times X \mid f^*(x^*, x) = \langle x^*, x \rangle \}
\]

\[
M_{\tilde{f}} := \{(x^*, x^*) \in X^* \times X^{**} \mid \tilde{f}^*(x^*, x^{**}) = \langle x^*, x^{**} \rangle \}.
\]

A now standard calculation [15, Prop. 4] shows that these sets are monotone when the associated functions are representative.

**Lemma 6** Suppose \( f : X \times X^* \rightarrow \mathbb{R} \) is a representative function. Suppose in addition that \( \tilde{f}^*(x^*, x^{**}) \geq \langle x^*, x^* \rangle \) (and hence \( f^*(x^*, x) \geq \langle x^*, x^* \rangle \)). Then

1. If \((u^*, u) \in X^* \times X \) is monotonically related to \( M_{\tilde{f}} \), we have \((u^*, u) \in M_{f^*} \).

2. The restrictions \( R_{X^* \times X} \left( M_{\tilde{f}} \right) = M_{f^*} \), \( R_{X \times X^*} \left( \widetilde{M}_{\tilde{f}} \right) = (M_{f^*})^\dagger \) for any \( \widetilde{M}_{\tilde{f}} \), a maximal monotone extension of \( \left( M_{\tilde{f}} \right)^\dagger \) (viewed as a monotone set within \( X^{**} \times X^* \)).

8
Proof. Let \((u^*, u) \in X^* \times X\) be monotonically related to \(M_{\hat{f}}\) (i.e., \(\langle u^* - y^*, u - y^* \rangle \geq 0 \) for all \((y^*, y^*) \in M_{\hat{f}}\)). As \(\Delta \geq 0\) we have \(f(z, z^*) - \langle z, z^* \rangle + \Delta (u - z, u^* - z^*) \geq 0\) for all \((z, z^*) \in X \times X^*\). Thus by Lemma 5 there exists \((x^*, x^*) \in X^* \times X^*\) such that

\[
\hat{f}^*(x^*, x^*) - \langle x^*, x^* \rangle + \Delta (u - x^*, u^* - x^*) \leq 0. \tag{10}
\]

It follows that \(\hat{f}^*(x^*, x^*) \leq \langle x^*, x^* \rangle\). Hence \(\hat{f}^*(x^*, x^*) = \langle x^*, x^* \rangle\) or \((x^*, x^*) \in M_{\hat{f}}\) and also by (10) we have

\[
\Delta (u - x^*, u^* - x^*) = \frac{1}{2} \lVert u - x^* \rVert^2 + \langle u - x^*, u^* - x^* \rangle + \frac{1}{2} \lVert u^* - x^* \rVert^2 = 0. \tag{11}
\]

Because \(\langle u - x^*, u^* - x^* \rangle \geq 0\) for all \((x^*, x^*) \in M_{\hat{f}}\) (as \((u, u^*)\) is monotonically related to \(M_{\hat{f}}\)) from (11) it follows that that \((u, u^*) = (x^*, x^*) \in M_{\hat{f}}\) and that \(x^*\) is actually in \(X\). That is \((u, u^*) \in R_{X \times X}(M_{\hat{f}})\) the restriction of \(M_{\hat{f}}\) into \(X^* \times X\).

On the other hand we know that the restriction of \(\hat{f}^*\) to \(X^* \times X\) is also a representative function (indeed that of \(M_{\hat{f}}\)) and so \(R_{X \times X}(M_{\hat{f}}) = M_{\hat{f}}\). Thus, if \((u^*, u) \in X^* \times X\) is monotonically related to \(M_{\hat{f}}\) then \((u, u^*) \in \hat{f}_{\hat{f}}\).

Clearly, when \((u^*, u) \in X^* \times X\) is monotonically related to \(M_{\hat{f}}\) then it is monotonically related to \(M_{\hat{f}}\) and so the same results follow for \(M_{\hat{f}}\).

An immediate corollary is the following.

**Corollary 7** Suppose \(f : X \times X^* \to \overline{R}\) is a representative function. Suppose in addition that \(\hat{f}^*(x^*, x^*) \geq \langle x^*, x^* \rangle\) (and hence \(\hat{f}^*(x^*, x) \geq \langle x, x^* \rangle\)). Let \(\hat{M}_{\hat{f}}\) be any maximal monotone extension of \(J_{X \times X^*}(M_{\hat{f}})\) (viewed as a subset of \(X^* \times X^*\)).

Then \(\hat{M}_{\hat{f}}\) is a maximal extension of \(M_{\hat{f}}\) with \(R_{X \times X^*}(\hat{M}_{\hat{f}}) = M_{\hat{f}}\).

**Proof.** Note that we have \(J_{X \times X^*}(M_{\hat{f}}) \subseteq \hat{M}_{\hat{f}}\) (for any extension as defined in Lemma 6) and so \(\hat{M}_{\hat{f}}\) is a maximal extension of the monotone set \(M_{\hat{f}}\). By Lemma 6 we have \(R_{X \times X^*}(\hat{M}_{\hat{f}}) = M_{\hat{f}}\).

An elegant result that holds in all Banach spaces is possible by using the composition result of [3] in conjunction with Theorem 8 below. This requires a qualification assumption and so in order to obtain the best result we delay statement of the qualification assumption to the next section.

The following theorem allows us to make the inference that the requisite Fitzpatrick function is indeed a representative function. This is often a crucial step toward proving the maximality of the underlying monotone set.

**Theorem 8** Suppose \(f : X \times X^* \to \overline{R}\) is a representative function with \(\hat{f}^*(x^*, x^*) \geq \langle x^*, x^* \rangle\) for all \((x^*, x^*) \in X^* \times X^*\). Suppose in addition that there exists an
family of subspaces $\mathcal{Y}$ of $X$ such that $\bigcup_{Y \in \mathcal{Y}} Y = X$ and a maximal monotone extension $T$ of the monotone set $J_{X \times X^*}(M_f^1) \subseteq X^{**} \times X^*$ with the property that whenever

$$A : Y \to X^{**}$$

is the embedding of a subspace $Y \in \mathcal{Y}$ into $X^{**}$ we have $T_A := A^* \circ \tilde{T} \circ A$ maximal monotone from $Y$ to $Y^*$.

Then when we have $\mathcal{F}_{M_f^*} : X \times X^* \to \mathbb{R}$ a representative function and $M_f = M_{f^*}$.

**Proof.** Note that as $\text{dom}(A^*) = X^{***}$ and range $\tilde{T} \subseteq X^*$ we should write $T_A := A^* \circ J_{X^{**}} \circ \tilde{T} \circ A$ but choose to accept the above simplification of notation. Let $Y \in \mathcal{Y}$ then for each $x \in Y$ we have $A(x) = J_X(x) = \hat{x}$ and when $z^* \in Y^*$ there exists $x^* \in X^*$ with $z^* = x^*|_Y$. Consequently when $(x, z^*) \in T_A$ then there exists $\hat{x} \in J_X(X)$ and $x^* \in \tilde{T}(\hat{x})$ such that $z^* = x^*|_Y$ and using Lemma 6 we have

$$\mathcal{F}_{\hat{M}_f^*}(\hat{x}, x^*) = \sup_{(y, y^*) \in (M_f^1)} \{ \langle y, x^* \rangle + \langle y^*, \hat{x} \rangle - \langle y, y^* \rangle \}$$

$$= \sup_{(y, y^*) \in \mathbb{R} \times Y^*} \{ \langle y, x^* \rangle + \langle y^*, \hat{x} \rangle - \langle y, y^* \rangle \}$$

$$\geq \sup_{(y, y^*) \in T_A \cap (Y \times X^*)} \{ \langle y, x^* \rangle + \langle y^*, \tilde{T}(\hat{x}) \rangle - \langle y, y^* \rangle \}$$

$$= \sup_{(y, y^*) \in T_A} \{ \langle y, z^* \rangle + \langle y^*, x \rangle - \langle y, y^* \rangle \} = \mathcal{F}_{T_A}(x, z^*)$$

(12)

As $T_A$ is maximal on $X \times Y^*$ we have for all $(\hat{x}, z^*) \in Y \times Y^*$ with $z^* = x^*|_Y$ that

$$\mathcal{F}_{T_A}(x, z^*) \geq \langle \hat{x}, z^* \rangle$$

and so

$$\mathcal{F}_{T_A}(x, z^*) = \sup_{(y, y^*) \in T_A} \{ \langle y, x^* \rangle + \langle y^*, x \rangle - \langle y, y^* \rangle \}$$

$$= \sup_{(y, y^*) \in T_A} \{ \langle y, x^* \rangle + \langle y^*, \hat{x} \rangle - \langle y, y^* \rangle \}$$

$$= \mathcal{F}_{T_A}(\hat{x}, x^*) \geq \langle \hat{x}, z^* \rangle = \langle \hat{x}, x^* \rangle$$

(13)

where $\mathcal{F}_{T_A}(\hat{x}, x^*)$ is the Fitzpatrick function of $T_A$ embedded in $X^{**} \times X^*$. As $J_{X \times X^*}(M_f^1) \subseteq (M_f^1)^\dagger \subseteq \tilde{T}$ we have on $X^{**} \times X^*$ pointwise,

$$\mathcal{F}_{M_f^*} \geq \mathcal{F}_{M_f^*}$$

Evaluating at $(\hat{x}, x^*)$ for $x \in Y$ we obtain from (12) and (13)

$$\mathcal{F}_{M_f^*}(x, x^*) = \mathcal{F}_{\hat{M}_f^*}(\hat{x}, x^*) \geq \langle \hat{x}, x^* \rangle.$$  

(14)

Since this holds for any choice of subspace $Y \in \mathcal{Y}$ we have (14) holding for all $x \in X$. Thus $\mathcal{F}_{M_f^*}$ is a representative function on $X \times X^*$.  

10
As \( f^*(x^*, \hat{x}) \geq \langle \hat{x}, x^* \rangle \) for all \((x^*, x) \in X^* \times X\) we have \( f^* \) is a representative function for \( M_{f^*} \). By [5, Prop. 2] \( \mathcal{P}_{M_{f^*}} (\cdot, \cdot) \geq f^*(\cdot, \cdot) \geq \mathcal{F}_{M_{f^*}} (\cdot, \cdot) \geq \langle \cdot, \cdot \rangle \).

Using the duality with respect to the pairing of \( X \times X^* \) with \( X^* \times X \), on taking conjugates we have

\[
\mathcal{P}_{M_{f^*}} = \mathcal{F}_{M_{f^*}} (\cdot, \cdot) \geq f^*(\cdot, \cdot) \geq \mathcal{F}_{M_{f^*}} (\cdot, \cdot) \geq \langle \cdot, \cdot \rangle.
\]

As \( \mathcal{F}_{M_f} (x, x^*) = \langle x, x^* \rangle \) by (15)

\[
M_f = \{ (x^*, x) | \mathcal{P}_{M_f} (x, x^*) = \langle x, x^* \rangle \} = \{ (x^*, x) | f (x, x^*) = \langle x, x^* \rangle \} = M_{f^*}.
\]

Also (15) implies

\[
M_f \subseteq \{ (x^*, x) | \mathcal{F}_{M_f} (x, x^*) = \langle x, x^* \rangle \}
\]

and [5, Prop. 2] says that \( \mathcal{F}_{M_f} (x, x^*) = \langle x, x^* \rangle \) implies \( \mathcal{P}_{M_f} (x, x^*) = \langle x, x^* \rangle \)

(or \((x, x^*) \in M_f \)) when \( \mathcal{F}_{M_f} \) is representative. So \( M_f \subseteq M_{f^*} \).

**Remark 9** When \( f \) is a representative function of a maximal monotone operator then we have already noted that \( M^\dagger_f \) also corresponds to the graph of this maximal monotone operator. As \( R_{X \times X^*} (M^\dagger_f) = M^\dagger_f \), we have \( T_A := A^* \circ M^\dagger_f: Y \rightarrow Y^* \).

We cannot deduce from the fact that \( \mathcal{F}_T \) is a representative function that \( T \) is maximal (graph closure may easily fail). In [5] a monotone operator \( T \) is said to be *almost maximal* when \( \mathcal{F}_T \) is a representative function, i.e., \( \mathcal{F}_T (x, x^*) \geq \langle x, x^* \rangle \). The close relationship between almost maximality and maximality for sums of maximal monotone operators is noted in [5]. The significance of Theorem 8 is that it gives sufficient conditions for almost maximality. We will use this to deduce maximality of a sum of two maximal monotone operators.

## 4 Properties of Operators in Finite Dimensions

We will be embedding operators into finite dimensional subspaces and will impose conditions to obtain maximality. Consequently the properties of finite dimensional maximal monotone operators are of some interest. In the ensuing analysis we need to reconstruct a finite dimensional maximal monotone operator from information provided only within the relative interior of its domain. This is considered in this section along with some related issues.

By \( \text{lin} C \) we denote the smallest linear subspace containing the set \( C \). We write \( 0 \in \text{sqri} S \) iff \( Z := \text{span} S \) is a closed subspace and \( 0 \in \text{core}_Z S \) (the core relative to \( Z \)). This has been referred to as the *strong quasi-relative interior condition* in the literature [13]. The presumption of a *strong quasi-relative interior* is generally a stronger assumption than the presumption of a *relative...
interior $\text{ri} \ S$, which corresponds to the interior relative to the affine hull of $S$. Indeed, if $S$ is closed $\text{sqri} \ S \subseteq \text{ri} \ S$ (see Theorem 3.6 of [13]).

An important piece of our technology is the following composition theorem first proved in [3]. If we take $A$ to denote the embedding of a finite dimensional subspace $Y$ into a nonreflexive Banach space $X$ we obtain a monotone operator $T_A := A^* \circ T \circ A : Y \to Y^*$ on a finite dimensional space $Y$.

**Theorem 10 ([3])** Let $T : X \to X^*$ is maximal monotone. Suppose in addition that $A : Y \to X$ is a continuous linear operator between two Banach spaces with $Y$ reflexive. Then $T_A := A^* \circ T \circ A$ is maximal on $Y$ when

$$0 \in \text{core} \ (\text{co dom} \ T + \text{range} \ A).$$

(16)

**Remark 11** Under assumption (16) one can show that

$$F^* (y^*, x^*) = \min \{ P_T (x^**, x^*) \mid A^* x^* = y^* \}.$$ for $F (y, x^*) = F_T (Ay, x^*)$. Hence

$$F^* (y^*, Ay) = \min \{ P_T (Ay, x^*) \mid A^* x^* = y^* \}$$

is a representative function of the maximal monotone set $T_A := A^* \circ T \circ A$, under the same assumptions as Theorem 10. Thus $F^* (y^*, Ay) = +\infty$ if $Ay \notin \text{Pr}_X \text{dom} \ P_T$.

This last observation holds more generally as we will shortly show.

**Lemma 12** Suppose $T$ is a maximal monotone operator on a Banach space $X$. Then

$$\text{Pr}_X \text{co} T = \text{co Pr}_X T = \text{co dom} T.$$ (17)

**Proof.** First note that $x \in \text{Pr}_X \text{co} T$ implies the existence of $x^*$ such that $(x, x^*) \in \text{co} T$ and hence the existence of $(x_i, x_i^*) \in T$ with $(x, x^*) = \sum \lambda_i (x_i, x_i^*) = (\sum \lambda_i x_i, \sum \lambda_i x_i^*)$. Thus $x = \sum \lambda_i x_i \in \text{co Pr}_X T$ and so $\text{Pr}_X \text{co} T \subseteq \text{co Pr}_X T$. Also $\text{Pr}_X \text{co} T \supseteq \text{co Pr}_X T$ since $\text{co Pr}_X T$ is the smallest convex set containing $\text{Pr}_X T$. That is, (17) holds. We will need to use the following partial conjugate formula (first noted by [15] in the case when $X$ is reflexive).

**Lemma 13** Suppose $p : X \times X^* \to \mathbb{R}$ is a proper lower semi–continuous function with $p^* = f$ where the conjugate is taken with respect to the paired spaces $X \times X^*$ and $X^* \times X$ with product topologies formed by endowing $X$ with the strong topology and $X^*$ with the bounded weak* topology. Let $h(u, v) := -(p(u, \cdot))^* (v)$ then for all $x, z \in X$ and $x^* \in X^*$ we have

$$f (x^*, x) = (h (\cdot, x))^* (x^*)$$ and

$$(f (x, \cdot))^* (z) = (h (\cdot, x))^* (z).$$ (18)
If we place $F(y, x^*) := f(Ay, x^*)$ where $A : Y \to X$ is a continuous linear mapping form the Banach space $Y$ into the the Banach space $X$ we have

$$F^*(y^*, x) = \sup_{z \in Y} \left\{ (z, y^*) + (h(\cdot, Az))^*(x) \right\}$$  \hspace{1cm} (19)

**Proof.** As $p^* = f$, performing this conjugation in steps let

$$-h(u, v) := (p(u, \cdot))^*(v)$$

$$= \sup_{z^*} \{ \langle v, z^* \rangle - p(u, z^*) \}$$  \hspace{1cm} (20)

and then

$$f(x, x^*) = p^*(x^*, x) = \sup_u \left( \langle u, x^* \rangle - (\sup_{z^*} \{ \langle x, z^* \rangle - p(u, z^*) \}) \right)$$

$$= \sup_u \left( \langle u, x^* \rangle - h(u, x) \right) = (h(\cdot, x))^*(x^*).$$

Consequently $(f(x, \cdot))^*(z) = (h(\cdot, x))^*(z)$. Now

$$F^*(y^*, x) = \sup_{(z, x^*) \in Y \times X^*} \left\{ (z, y^*) + \langle x, x^* \rangle - f(Az, x^*) \right\}$$

$$= \sup_{z \in Y} \left\{ (z, y^*) + \sup_{x^* \in X^*} \{ \langle x, x^* \rangle - f(Az, x^*) \} \right\}$$

$$= \sup_{z \in Y} \left\{ (z, y^*) + (f(Az, \cdot))^*(x) \right\} = \sup_{z \in Y} \{ (z, y^*) + (h(\cdot, Az))^*(x) \}. $$

A characterisation of the domain of $F^*$ may now be established.

**Lemma 14** Let $A : Y \to X$ be a continuous linear operator between Banach spaces. Let $F_T$ be the Fitzpatrick function of a monotone operator $T : X \to X^*$ and denote $F_T(y, x) = F_T(Ay, x^*) : Y \times X^* \to \mathbb{R}$. Then $F^*(y^*, x) = +\infty$ if $x \notin \text{Pr}_X \text{ dom} T$. In particular $F^*(y^*, Ay) = +\infty$ when $Ay \notin \text{co dom} T$.

**Proof.** First note that $\text{Pr}_T = F_T$. Lemma 13 with $f = F_T$ and $p = \text{Pr}_T$ shows

$$(F_T(x, \cdot))^*(z) = (h(\cdot, x))^*(z)$$

where the later function corresponds the lower semi–continuous hull of the convex function $u \mapsto h(u, v) = -(\text{Pr}_T(u, \cdot))^*(v)$. Now using (19) we have

$$F^*(y^*, x) = \sup_{z \in Y} \{ (z, y^*) + (h(\cdot, Az))^*(x) \} \geq \langle y, y^* \rangle + (h(\cdot, Ay))^*(x).$$

Next note that

$$u \mapsto h(u, Ay) = -(\text{Pr}_T(u, \cdot))^*(Ay) = -\sup_{w^*} \{ \langle w^*, Ay \rangle - \text{Pr}_T(u, w^*) \}$$

$$= \inf_{w^*} \{ \text{Pr}_T(u, w^*) - \langle w^*, Ay \rangle \}. $$
When $x \notin \text{Pr}_X(\text{dom} P_T)$ there is $\delta > 0$ such that $B_\delta (x) \cap \text{Pr}_X(\text{dom} P_T) = \emptyset$ implying

\[(h(\cdot, Ay))^{**}(x) \geq \inf_{u \in B_\delta (x)} \inf_{w^*} \{P_T (u, w^*) - \langle w^*, Ay \rangle \} = +\infty.\]

When $T$ is maximal we now need to investigate the conjugate function of $F_T(Ay, x^*)$, as this will be the function we use in many results. It is closely related to a representative function for $T_A$.

**Lemma 15** Let $A : Y \to X$ be a continuous linear operator between Banach spaces. Let $F_T$ be the Fitzpatrick function of a maximal monotone operator $T : X \to X^*$ and denote $F_T(Ay, x^*)$. Then for all $(y^*, y) \in Y^* \times Y$ we have

\[F^* (y^*, Ay) \geq F_{A^* \circ T \circ A} (y^*, y) = \sup_{(z, z^*) \in T \circ A} \{(z^*, Ay) + \langle z, y^* \rangle - \langle z^*, Az \rangle \}.\]

(21)

**Proof.** By direct calculation

\[F^* (y^*, Ay) = \sup_{(z, z^*) \in Y \times X^*} \{(z^*, Ay) + \langle z, y^* \rangle - F (z, z^*) \} \geq \sup_{(z, z^*) \in T \circ A} \{(z^*, Ay) + \langle z, y^* \rangle - F_T (Az, z^*) \} = \sup_{(z, z^*) \in T \circ A} \{(z^*, Ay) + \langle z, y^* \rangle - \langle Az, z^* \rangle \}.

Also by direct computation

\[F_{A^* \circ T \circ A} (y^*, y) = \sup_{(w^*, w) \in A^* \circ T \circ A} \{(w^*, y) + \langle w, y^* \rangle - \langle w, w^* \rangle \} = \sup_{(x^*, x) \in T \circ A} \{(A^* x^*, y) + \langle w, y^* \rangle - \langle w, A^* x^* \rangle \} = \sup_{(x^*, x) \in T \circ A} \{(x^*, Ay) + \langle w, y^* \rangle - \langle Aw, x^* \rangle \}

and so (21) holds.

**Corollary 16** Let $A : Y \to X$ be a continuous linear operator between Banach spaces. Let $A^* : X^* \to Y^*$. Let $F_T$ be the Fitzpatrick function of a maximal monotone operator $T : X \to X^*$ and denote $F_T(Ay, x^*)$. Suppose in addition $A^* \circ T \circ A$ is maximal on $Y \times Y^*$. Then for all $(y^*, y) \in Y^* \times Y$ we have

\[F_{A^* \circ T \circ A} (y^*, y) \geq \langle y, y^* \rangle \quad \text{and} \quad F_{A^* \circ T \circ A} (y^*, y) = \langle y, y^* \rangle \iff y^* \in (A^* \circ T \circ A) (y).\]
Proof. This follows from the usual properties of the Fitzpatrick function of a maximal monotone set.

Let \( 0^+ A := \{ z \mid x + \lambda z \in A, \forall x \in A \text{ and } \lambda \geq 0 \} \) denote the recession directions of a convex set \( C \). We denote the normal cone to a close convex set \( C \) at a point \( x \in C \) by

\[
N_C(x) := \{ x^* \in X^* \mid \langle y - x, x^* \rangle \leq 0 \text{ for all } y \in C \}.
\]

The following result is probably well known and is easily proved.

**Lemma 17** Suppose \( M \) is a maximal monotone operator on a Banach space \( X \). Let \( C = \text{co} \text{ dom} \ M \) and \( x \in \text{dom} \ M \). Then

\[
0^+ M(x) = N_C(x).
\]  

(22)

In Proposition 25 of the next section, we define a maximal monotone operator \( Q_A \) on a finite-dimensional space \( Y \) induced by two maximal monotone operators \( T \) and \( M \) on a nonreflexive space \( X \). It will be shown that \( Q_A \) can only possible differ from \( (T + M)_A \) at boundary points of its domain. The next two results resolve the issue as to equality on the whole domain and finally allow the demonstration of our main sum theorem. In essence, we show \( Q_A \) can be reconstructed from the relative interior of its domain along with the addition of certain normal cones.

**Proposition 18** Suppose \( Q : Y \rightrightarrows Y^* \) is a maximal monotone operator with \( \text{int dom} \ Q \neq \emptyset \) and suppose \( Y \) is finite-dimensional. Let \( K(y) := Q(y) \) for all \( y \in \text{int dom} \ Q \) and for each \( y \in (\text{dom} \ Q) \cap (\text{int dom} \ Q)^c \) define

\[
K(y) := \text{co} \left( \limsup_{y' \in \text{int dom} \ Q \atop y' \to y} Q(y') \right) + N_{\text{dom} \ Q}(y).
\]

Then \( K \) is maximal monotone and coincides with \( Q \).

**Proof.** We make the usual interpretation of \( A + B = \emptyset \) if either \( A = \emptyset \) or \( B = \emptyset \). As \( Q \) is maximal we know that \( \text{int dom} \ Q \) is a convex set [3]. Suppose \( (y_0, y_0^*) \) is monotonically related to \( Q \). First note by [5, Cor. 16] that \( Q \) is maximal monotone locally. Thus when \( y_0 \in \text{int dom} \ Q \) we may choose a neighbourhood \( B_\delta(y_0) \subseteq \text{int dom} \ Q \) and hence, as \( (y_0, y_0^*) \) is monotonically related to \( Q|_{B_\delta(y_0)} \) we must have \( y_0^* \in Q(y_0) = K(y_0) \).

This leaves the situation where \( y_0 \in (\text{dom} \ Q) \cap (\text{int dom} \ Q)^c \). Suppose first that

\[
\limsup_{y \in \text{int dom} \ Q \atop y \to x} Q(y) \neq \emptyset.
\]

Then, there is \( \bar{r} > 0 \) such that \( Q(y) \cap \overline{B_r(0)} \neq \emptyset \) for all \( r \geq \bar{r} \), and the following (standard) argument suffices. As \( Y \) is finite-dimensional and \( K(y_0) \) is closed, by the Separation theorem and \( y_0^* \notin K(y_0) \) there are \( u \in X \) and \( \alpha \) with

\[
\langle y^*, u \rangle < \alpha < \langle y_0^*, u \rangle, \text{ for all } y^* \in K(y_0).
\]
As \( N_{\text{dom } Q}(y_0) \subseteq K(y_0) \) it follows that \( \langle n^*, u \rangle \leq 0 \) for all \( n^* \in N_{\text{dom } Q}(y_0) \) and so \( u \in T_{\text{dom } Q}(y_0) \) (the tangent cone at \( y_0 \)). Let \( \alpha < \beta < \langle y_0^*, u \rangle \).

Let \( W := \{ x^* : \langle x^*, u \rangle < \alpha \} \) be an open set containing \( K(y_0) \). By definition of \( K \) (and the coincidence of Kuratowski-Painlevé and epi-distance convergence within finite dimensions), for each \( r > \bar{r} \) there exists a neighbourhood \( B_r(y_0) \) such that \( K(y) \cap B_r(0) \subseteq W \) for all \( y \in B_r(y_0) \). Since \( u \in T_{\text{dom } Q}(y_0) \) there exists a \( t > 0 \) and \( v \in V := \{ x : \langle y_0^*, x \rangle > \beta \} \) such that \( \|u - v\| < (\beta - \alpha)/r \) and \( y + tv \in \text{dom } Q \cap B_r(y_0) \). Then \( K(y + tv) \cap B_r(0) \subseteq W \) and for all \( u^* \in K(y + tv) \cap B_r(0) \subseteq W \) we have

\[
0 \leq \langle y_0^* - u^*, y - y - tv \rangle = -t\langle y_0^* - u^*, v \rangle \quad \text{implying} \quad \langle y_0^*, v \rangle \leq \langle u^*, v \rangle.
\]

Thus

\[
\langle u^*, u \rangle = \langle u^*, v \rangle + \langle u^*, u - v \rangle \geq \langle y_0^*, v \rangle - \|u^*\|\|u - v\| \geq \beta - r(\beta - \alpha)/r = \alpha
\]

implying \( u^* \notin W \), a contradiction. Thus, \( y_0^* \in K(y_0) \).

This leaves only the case when \( K(y_0) = \emptyset \). We shall show that \( y_0 \notin \text{dom } Q \).

Once this is shown we have \( K|_{\text{dom } Q} \) maximal relative to \( \text{dom } Q \) and as we clearly have \( K \subseteq Q \) it follows that \( K(x) = Q(x) \) for all \( x \in \text{dom } Q \). The maximality of \( Q \) relative to \( Y \) implies \( K \) is maximal relative to \( Y \) as well. To this end assume \( y_0 \in \text{dom } Q \) (and hence \( y_0 \in \text{dom } Q \cap (\text{int dom } Q)^c \)), take \( y_0^* \in Q(y_0) \) and a sequence \( y_n \in \text{int } Y \cap \text{dom } Q \) such that \( \frac{y_n - y_0}{\|y_n - y_0\|} \to z \in \text{int } T_{\text{dom } Q}(y_0) \) (the latter being nonempty due to the assumption that \( \text{int dom } Q \neq \emptyset \)). Then by the monotonicity of \( Q \) we have for all \( n \),

\[
\langle y_n - y_0, y_n^* - y_0^* \rangle \geq 0 \quad \text{implying} \quad \langle \frac{y_n - y_0}{\|y_n - y_0\|}, \frac{y_n^* - y_0^*}{\|y_n^* - y_0^*\|} \rangle \geq 0.
\]

On taking subsequences we may assume \( \frac{y_n - y_0}{\|y_n - y_0\|} \to z^* \) and hence

\[
\langle z, z^* \rangle \geq 0 \quad \text{for some} \quad z \in \text{int } T_{\text{dom } Q}(y_0).
\]

On the other hand by maximality of \( Q \)

\[
\langle y - y_n, y_n^* - y_0^* \rangle \geq 0 \quad \text{for all} \quad (y, y^*) \in Q
\]

implying

\[
\lim_n \langle y - y_n, \frac{y_n^* - y_0^*}{\|y_n^* - y_0^*\|} \rangle \geq 0 \quad \text{or}
\]

\[
\langle y - y_0, -z^* \rangle \geq 0 \quad \text{for all} \quad y \in \text{dom } Q.
\]

Consequently, \( z^* \in N_{\text{dom } Q}(y_0) \) contradicting (23).

As we are in finite dimensions we may recast the last result.

**Corollary 19** Suppose \( Q : Y \rightharpoonup Y^* \) is a maximal monotone operator and that \( Y \) is finite-dimensional. Let \( K(y) := Q(y) \) for all \( y \in \text{ri dom } Q \) and for each \( y \in \overline{\text{dom } Q} \cap (\text{ri dom } Q)^c \) define

\[
K(y) := \text{co} \left( \limsup_{y' \to y} Q(y') \right) + N_{\text{dom } Q}(y) .
\]

16
Then $K = Q$.

**Proof.** Without loss of generality we may assume $0 \in \text{dom } Q$. Let $Z := \text{span } \text{dom } Q$. Then relative to $Z$ we have $\text{int}_Z \text{dom } Q \neq \emptyset$ (and $\text{int}_Z \text{dom } Q$ is convex due to maximality of $(\text{dom } Q)$).

Let $A : Z \to Y$ be the embedding of $Z$ into $Y$ and note that $\text{ri } (\text{Pr}_X \text{dom } F_Q) \cap \text{Range } A \neq \emptyset$ so we may apply Theorem 10 to obtain maximality of $Q_A := A^* \circ Q \circ A$. Apply Proposition 18 to $Q_A$ to obtain maximality of $\tilde{K}$ (as defined on $Z$). We have $\tilde{K} = Q \cap (Z \times Z^*)$, and we note, since $\text{dom } Q \subseteq Z$, that $\tilde{K}$ defined in (25) only differs from $\tilde{K}$ by the addition of the annihilator $Z^\perp$. Hence, $K$ extends $\tilde{K}$ from $Z$ to $Y$ as a maximal monotone operator. Consequently, $Q \subseteq K$ and the maximality of $Q$ implies $Q = K$. \hfill $\blacksquare$

We finish this section by studying the relationship between the normal cones to $\text{dom } M$ and to $(\text{dom } M) \cap Y$, since the domain of $M_A$ is equal to $(\text{dom } M) \cap Y$ for any maximal monotone operator $M : X \to X^*$.

**Lemma 20** Suppose $M$ is a maximal monotone operator on a Banach space $X$, and $Y$ is a finite dimensional subspace of $X$. Let $\bar{y} \in \text{bd } (\text{dom } M \cap Y)$. Then

$$M_A (\bar{y}) = M_A (\bar{y}) + (N_{\text{dom } M})_A (\bar{y}),$$

and

$$N_{(\text{dom } M) \cap Y} (\bar{y}) = (N_{\text{dom } M})_A (\bar{y}),$$

where $(N_{\text{dom } M})_A (\bar{y}) := (A^* \circ N_{\text{dom } M} \circ A) (\bar{y})$.

**Proof.** Maximality of $M$ implies that $M = M + N_{\text{dom } M}$, since the latter is a (larger) monotone operator. Consequently,

$$M_A (\bar{y}) = (A^* \circ M \circ A) (\bar{y}) = (A^* \circ M \circ A) (\bar{y}) + (A^* \circ N_{\text{dom } M} \circ A) (\bar{y}) = M_A (\bar{y}) + (N_{\text{dom } M})_A (\bar{y}).$$

Observe that both cones in (26) reside in $Y^*$ and that $\bar{y} \in Y$, giving $A\bar{y} = \bar{y}$. When $y^* \in N_{(\text{dom } M) \cap Y} (\bar{y})$ we have

$$\langle y^*, x - \bar{y} \rangle \leq 0 \quad \text{for all} \quad x \in (\text{dom } M) \cap Y.$$

As $Y$ is finite-dimensional it is complemented in $X$. We may thus write $X = Y \oplus Z$ with $Z$ closed. Extend $y^*$ to $x^* \in X^*$ via linearity by placing $\langle x^*, x \rangle = 0$ for $x \in Z$ and $\langle x^*, y \rangle = \langle y^*, y \rangle$ for $y \in Y$. Then we have

$$\langle x^*, x - \bar{y} \rangle \leq 0 \quad \text{for all} \quad x \in \text{dom } M.$$ 

That is $x^* \in N_{\text{dom } M} (\bar{y})$ with $x^*|_Y = y^*$ and we have

$$y^* = A^* x^* \in (A^* \circ (N_{\text{dom } M}) \circ A) (\bar{y}) = (N_{\text{dom } M})_A (\bar{y}),$$

establishing $N_{(\text{dom } M) \cap Y} (\bar{y}) \subseteq (N_{\text{dom } M})_A (\bar{y})$. 

17
Now take $y^* \in (N_{\text{dom } M})_A(y)$. Then there exists $x^* \in N_{\text{dom } M}(Ay)$ with $y^* = A^*x^*$ and so $\langle x - Ay, x^* \rangle \leq 0$ for all $x \in \text{dom } M$. Thus we also have $\langle x - Ay, x^* \rangle \leq 0$ for all $x \in \text{dom } M$ and consequently for all $x = Ay' \in \text{dom } M \cap \hat{Y}$ we have

$$\langle Ay' - Ay, x^* \rangle = \langle y' - y, A^*x^* \rangle = \langle y' - y, y^* \rangle \leq 0$$

for all $y' \in \text{dom } M \cap \hat{Y}$.

Thus, $\langle y' - y, y^* \rangle \leq 0$ for all $y' \in \text{dom } M \cap \hat{Y}$ or $y^* \in N_{\text{dom } M \cap \hat{Y}}(y)$. □

5 Conditions for Maximality of a Sum

There are important outstanding problems in nonreflexive spaces regarding the type of qualification that ensures maximality of a sum of maximal monotone operators. The following analysis is motivated by the desire to capture a generalization of the classic result of Rockafellar [19].

In [16] the following was observed. Let $f, g : X \times X^* \to \mathcal{R}$ be proper, closed convex functions and

$$(f \square_2 g)(x, x^*) := \inf \{g(x, y_1^*) + h(x, y_2^*) \mid x^* = y_1^* + y_2^*, y_1^*, y_2^* \in X^*\}.$$  

Then

$$(f \square_2 g)^*(x^*, x^{***}) \leq (f \square_1 g^*)(x^*, x^{***})$$

$$:= \inf \{f^*(x_1^*, x^{***}) + g^*(x_2^*, x^{**}) \mid x_1^* + x_2^{**} = x^*, x_1^*, x_2^{**} \in X^*\}. \quad (27)$$

In [16] the problem of proving $T + M$ is maximal in a reflexive space is reduced to questions regarding equality in (27) and the nature of the conjugates $f^*$ and $g^*$. When we study $f = \mathcal{F}_M$ for a maximal monotone operator then by Theorem 3 we have $\hat{f}^*(x^*, x^{**}) = \hat{\mathcal{F}}_M(x^*, x^{**}) \geq \langle x^{**}, x^* \rangle$ on $X^{**} \times X^*$. Let $\hat{x} = J_X(x) \in X^{**}$ and recall that such inequalities imply corresponding ones in the embedded space $X \times X^*$ in that $\hat{f}^*(x^*, x) = \mathcal{P}_M(x, x^*) = \hat{\mathcal{F}}_M(x^*, \hat{x}) \geq \langle x, x^* \rangle$ on $X \times X^*$. The following is a slight generalization of a result in [28].

**Theorem 21** Suppose $T$ and $M$ are maximal monotone operators and $f$ and $g$ are their respective representative functions. Suppose that $f$ and $g$ may be chosen so that equality holds in (27) for all $(x^*, x^{**}) \in X^* \times X^{**}$ with the infimum being attained in the definition of $(f \square_1 g^*)^*(x^*, x^{**})$ for some $x_1^*, x_2^* \in X^*$.

Then both $h := (f \square_2 g)$ and $h^* = (f^* \square_1 g^*)$ are representative functions with $(h^*)^\dagger$ a representative function for $T + M$.

If $f$ and $g$ are chosen so that $\hat{f}^*(\cdot, \cdot) \geq \langle \cdot, \cdot \rangle$ and $\hat{g}^*(\cdot, \cdot) \geq \langle \cdot, \cdot \rangle$ then $\hat{h}^*(\cdot, \cdot) \geq \langle \cdot, \cdot \rangle$ on $X^* \times X^{**}$.

**Proof.** Let $h := (f \square_2 g)$. First note that as $f(x, y_1^*) \geq \langle x, y_1^* \rangle$ and $g(x, y_2^*) \geq \langle x, y_2^* \rangle$ we have

$$h(x, x^*) = \inf \{f(x, y_1^*) + g(x, y_2^*) \mid x^* = y_1^* + y_2^*, y_1^*, y_2^* \in X^*\}$$

$$\geq \inf \{\langle x, y_1^* \rangle + \langle x, y_2^* \rangle \mid x^* = y_1^* + y_2^*, y_1^*, y_2^* \in X^*\} = \langle x, x^* \rangle$$

18
and so \( f \) is a representative function. Since \( T \) and \( M \) are monotone \( \hat{f}^* (x_1^*, x^{**}) \geq \langle x_1^*, x^{**} \rangle \) and similarly \( \hat{g}^* (x_2^*, x^{**}) \geq \langle x_2^*, x^{**} \rangle \) (as noted in Corollary 3). A similar calculation shows

\[
\hat{h}^* (x^*, x^{**}) = \left( \hat{f}^* \square_1 \hat{g}^* \right) (x^*, x^{**}) \geq \langle x^*, x^{**} \rangle.
\]

Because \( g \) and \( h \) are representative for their respective maximal monotone sets we have

\[
M_f := \{(x, x^*) \in X \times X^* \mid f (x, x^*) = \langle x, x^* \rangle \} = \{(x, x^*) \in X \times X^* \mid f^* (x^*, x) = \langle x^*, x \rangle \} = M_{f^*} = T.
\]

Similarly, \( M_g = M_{g^*} = M \). Let \( (x, x^*) \in M_{h^*} \). By assumption, there exists \( x_1^*, x_2^* \in X^* \) with \( x_1^* + x_2^* = x^* \) such that

\[
h^* (x^*, x) = (f^* \square_1 g^*) (x^*, x) := f^* (x_1^*, x) + g^* (x_2^*, x) = \langle x^*, x \rangle = \langle x_1^*, x \rangle + \langle x_2^*, x \rangle.
\]

As \( f^* (x_1^*, x) \geq \langle x_1^*, x \rangle \) similarly \( g^* (x_2^*, x) \geq \langle x_2^*, x \rangle \) and

\[
f^* (x_1^*, x) - \langle x_1^*, x \rangle + g^* (x_2^*, x) - \langle x_1^*, x \rangle = 0
\]

we have \( f^* (x_1^*, x) = \langle x_1^*, x \rangle \) and \( g^* (x_2^*, x) = \langle x_1^*, x \rangle \). Hence \( (x, x_1^*) \in M_{f^*} = T \) and \( (x, x_2^*) \in M_{g^*} = M \) with \( x^* = x_1^* + x_2^* \in T(x) + M(x) \). Thus \( (M_{h^*})^1 \) is contained in the graph of \( T + M \) which is itself a monotone set. Suppose now that \( (x, x_1^*) \in M_{f^*} = T \) and \( (x, x_2^*) \in M_{g^*} = M \) then as \( f^* (x_1^*, x) = \langle x_1^*, x \rangle \) and \( g^* (x_2^*, x) = \langle x_2^*, x \rangle \) we have

\[
h^* (x^*, x) = (f^* \square_1 g^*) (x^*, x) \leq f^* (x_1^*, x) + g^* (x_2^*, x) = \langle x_1^*, x \rangle + \langle x_2^*, x \rangle = \langle x^*, x \rangle
\]

implying \( h^* (x^*, x) = \langle x^*, x \rangle \). Thus, \( (h^*)^1 \) is a representative function for the sum \( T + M \).

It is shown in a Banach space in [26] that for \( g, h \) proper, lower semi-continuous convex functions we have

\[
\left( f \square_2 g \right)^* (x^*, y^*) = \left( f^* \square_1 g^* \right) (x^*, y^*)
\]

under the assumption that

\[
0 \in \text{sqri} (\text{Pr}_X (\text{dom} f) - \text{Pr}_X (\text{dom} g))
\]

where \( \text{Pr}_X \) is the projection onto \( X \) and we write \( 0 \in \text{sqri} S \) iff \( Z := \text{span} S \) is a closed subspace and \( 0 \in \text{core}_Z S \) (the core relative to \( Z \)).

We next note that \( f \square_2 g \), produces a representative function for \( T + M \) under another weak qualification assumption. Recall, that to show a convex set is bounded weak* closed it is sufficient to show it contains the limits of all bounded and weak* convergent nets taken from the set [14].
Theorem 22 Suppose \( T \) and \( M \) are maximal monotone operators from a Banach space \( X \) to \( X^* \). Suppose in addition there exist representative functions \( g \) and \( h \) for \( T \) and \( M \) respectively such that

\[
0 \in \text{core} \left( \text{Pr}_X (\text{dom} f^*) - \text{Pr}_X (\text{dom} g^*) \right)
\]

Then \( h := (f \Box g) (x, x^*) \) is proper, \( s \times \text{bw}^* \)-closed. Moreover, \( h \) is exact and also is a representative function for \( T + M \).

**Proof.** For each \( x \in X \) and \( K > \inf_{x^* \in X^*} (f \Box g) (x, x^*) \) let

\[
H (K, M, r) := \left\{ (x_1^*, x_2^*) \in X^* \times X^* \mid f (x, x_1^*) + g (x, x_2^*) \leq K \right\}
\]

with \( \|x_1^* + x_2^*\| \leq r \) and \( \|x\| \leq M \).

We claim that there exists \( C (K, M, r, u, v) \) such that for all \((x^*, y^*) \in H (K, M, r)\) we have \((x^*, u) + (y^*, v) \leq C (K, M, u, v)\).

Indeed, since (29) holds \((u - v) = \lambda (a^* - b^*) \) for some \( \lambda > 0 \) and some \((a^*, a) \in \text{dom} g^* \) and \((b^*, b) \in \text{dom} h^*\). Thus \( u - v = \lambda (a - b) \). Then using the Fenchel inequality

\[
\langle x^*, u \rangle + \langle y^*, v \rangle = \lambda \langle x^*, a \rangle + \lambda \langle y^*, b \rangle + \langle x^* + y^*, v - \lambda b \rangle
\]

\[
= \lambda \langle (x^*, a) + (x^*, a^*) \rangle + \lambda \langle (y^*, b) + \langle x^*, b^* \rangle \rangle
\]

\[
+ \langle x^* + y^*, v - \lambda b \rangle - \lambda \langle x, a^* + b^* \rangle
\]

\[
\leq \lambda (f^* (a^*, a) + f (x, x^*) + g^* (b^*, b) + g (x, y^*))
\]

\[
+ \|x^* + y^*\| \|v - \lambda b\| + \lambda M \|a^* + b^*\|
\]

\[
\leq \lambda (K + f^* (a^*, a) + g^* (b^*, b) + M \|a^* + b^*\|)
\]

\[
+ r \|v - \lambda b\| := C (K, M, r, u, v).
\]

Note that this bound only depends on \( x \) in so far as we require that the inequalities \( \|x\| \leq M \) and \( K \geq (f \Box g) (x, x^*) \) are satisfied.

Thus, by the uniform boundedness principle we have \( H (K, M, r) \) contained in a ball whose radius depends only on the choice of \( K, M \) and \( r \). As \( H (K, M, r) \) is clearly closed it is weak* compact. We now show that the level sets of \((f \Box g)\) are \( s \times \text{bw}^* \)-closed. Take \( (x_\beta, x^*_\beta) \in \{ (x, x^*) \mid (f \Box g) (x, x^*) \leq K \} \) such that \( \{x^*_\beta\} \) is a bounded weak* convergent net converging to \( x^* \) and \( x^*_\beta \) strongly converging to \( x \). Let \( \varepsilon_\beta > 0 \) with \( \varepsilon_\beta \downarrow 0 \) and \( r > 0 \) such that \( \max \{ \varepsilon_\beta, \|x^*_\beta\| \} \leq r \) along with \( \|x_\beta\| \leq 2 \|x\| := M \)

\[
(u^*_\beta, v^*_\beta) \in H (K + \varepsilon_\beta, M, r) \quad \text{with} \quad u^*_\beta + v^*_\beta = x^*_\beta
\]

and by the previous observation we have \( \left\| (u^*_\beta, v^*_\beta) \right\| \) uniformly bounded for all \( \beta \). Thus on passing to a subnet, we may assume \( (u^*_\beta, v^*_\beta) \overset{w^*}{\to} (u^*, v^*) \) for some
$u^*, v^* \in X^*$ with $u^* + v^* = x^*$. Then, since $x_\beta^* \to x^*$, we conclude

$$(g \Box_2 h)(x, x^*) \leq f(x, x^* - v^*) + g(x, v^*)$$

$$\leq \liminf_\beta (f(x, x^* - v^*) + g(x, v^*_\beta))$$

$$\leq \liminf_\beta \left( f(x, x^* - v^*_\beta) + g(x, v^*_\beta) \right) \leq \liminf_\beta (K + \varepsilon_\beta) = K.$$ 

Thus the level-set of $f \Box_2 g$ is s- $bw^*$-closed and so $(x, x^*) \mapsto (f \Box_2 g) (x, x^*)$ is $s \times bw^*$- lower semi-continuous.

To verify properness, suppose to the contrary that $(f \Box_2 g)(x^*) = -\infty$ for some $(x, x^*)$. Then $\lim_\beta (f(x, x^* - v^*_\beta) + g(x, v^*_\beta)) = -\infty$ for some net $v^*_\beta$. Without loss of generality we may assume $\lim_\beta (f(x, x^* - v^*_\beta) + g(x, v^*_\beta)) \leq 0$ for all $\beta$ and since $(x^* - v^*_\beta, v^*_\beta) \in H(0, \|x^*\|, x_\beta)$ with this set being uniformly bounded, we have $v^*_\beta \to v^*$ on taking a subnet. Then

$$-\infty < f(x, x^* - v^*) + g(x, v^*) \leq \liminf_\beta (f(x, x^* - v^*_\beta) + g(x, v^*_\beta))$$

$$\leq \liminf_\beta \left( f(x, x^* - v^*_\beta) + g(x, v^*_\beta) \right) = -\infty,$$

a contradiction, and so $(f \Box_2 g)$ is proper.

Next, to show $(f \Box_2 g) (x, x^*)$ is exact, we take $K = (f \Box_2 g) (x, x^*)$ and $\varepsilon_\beta > 0$ with $\varepsilon_\beta \downarrow 0$. Then there exists

$$(u^*_\beta, v^*_\beta) \in H(K + \varepsilon_\beta, \|x\|, x) \quad \text{with} \quad u^*_\beta + v^*_\beta = x^*.$$ 

As $\|(u^*_\beta, v^*_\beta)\|$ is bounded, on taking $w^*$-convergent subnet we may assume

$$(u^*_\beta, v^*_\beta) \to (u^*, v^*) \quad \text{with} \quad u^* + v^* = x^*$$

and so

$$(f \Box_2 g)(x, x^*) = f(x, x^* - v^*) + g(x, v^*),$$

giving the exactness of $(f \Box_2 g)$.

Finally

$$\{(x, x^*) \mid (f \Box_2 g)(x, x^*) = (x, x^*)\}$$

$$= \{(x, x^*) \mid \exists v^* \in X^* \text{ such that } f(x, x^* - v^*) + g(x, v^*) = \langle x, x^* \rangle\}$$

$$= \{(x, x^*) \mid \exists v^* \in X^* \text{ such that } f(x, x^* - v^*) - \langle x, x^* - v^* \rangle + g(x, v^*) = \langle x, v^* \rangle = 0\}$$

$$= \{(x, x^*) \mid \exists v^* \in X^* \text{ such that } f(x, x^* - v^*) = \langle x, x^* - v^* \rangle \text{ and } g(x, v^*) = \langle x, v^* \rangle\}$$

$$= \{(x, x^*) \mid \exists v^* \in X^* \text{ such that } x^* - v^* \in T(x) \text{ and } v^* \in M(x)\}$$

$$= \{(x, x^*) \mid x^* \in (T + M)(x)\}.$$ 

The previous result gives us a way of deducing that $(g \Box_2 h)$ is a representative function for $T + M$ which is not covered directly by Theorem 21 under the assumption of (29). In order to obtain a sum theorem for two maximal
monotone operators we need the following chain of results that investigates
the particular representative function \( H (y, y^*) := \inf \{ h (Ay, x^*) \mid A^* x^* = y^* \} \)
(with \( h := (f \Box g) \)) and its conjugate. Note that under the qualification assumption (29) \( h \) is a representative function for \( T + M \), and so for all \( (y, y^*) \) we have

\[
H (y, y^*) = \inf \{ h (Ay, x^*) \mid A^* x^* = y^* \} \\
\geq (Ay, x^*) = \langle y, A^* x^* \rangle = \langle y, y^* \rangle. \tag{30}
\]

Consequently \( H \) may be used to define a monotone set \( Q_A \). Note also that in order that \( y \in \Pr_X \dom H \) there must exist \( x^* \) such that \(+\infty > h (Ay, x^*)\) implying

\[
Ay \in \Pr_X \dom h \quad \text{or} \quad y \in \Pr_X \dom h \cap \hat{Y}.
\]

Hence under the assumption of (29) and the exactness of the infimal convolution

\[
\Pr_X \dom H = \Pr_X \dom h \cap \hat{Y} = \Pr_X \dom f \cap \Pr_X \dom g \cap \hat{Y}. \tag{31}
\]

Finally note that when the qualification assumption (29) holds, as \( h \) is a representative function for \( T + M \), for \( y^* \in (T + M) (Ay) \) with \( y^* = A^* x^* \) we have

\[
H (y, y^*) = \inf \{ h (Ay, x^*) \mid A^* x^* = y^* \} \\
= (Ay, x^*) = \langle y, A^* x^* \rangle = \langle y, y^* \rangle
\]

and so \( (y, y^*) \in Q_A \). Consequently we always have

\[
(T + M)_A \subseteq Q_A. \tag{32}
\]

**Proposition 23** Let \( A : Y \to X \) be a continuous linear operator between Banach spaces. Let \( f, g : X \times X^* \to \overline{\mathbb{R}} \) be proper, lower semi-continuous with

\[
\emptyset \neq \Range A \cap \left[ \core \Pr_X (\dom f) \right] \cap \Pr_X (\dom g). \tag{33}
\]

Then for \( h := (f \Box g) \) and \( H (y, y^*) := \inf \{ h (Ay, x^*) \mid A^* x^* = y^* \} \) we have

\[
H^* (y^*, y) = (F^* \Box G^*) (y^*, Ay)
\]

for all \((y^*, y) \in Y^* \times Y\), where \( F (y, x^*) = f (Ay, x^*) \) and \( G (y, x^*) = g (Ay, x^*) \).

Moreover, the infimal convolution is exact.

**Proof.** We have

\[
H^* (v^*, v) = \sup_{(y,y^*)} \left\{ \langle y, v^* \rangle + \langle y^*, v \rangle - \inf_x \{ h (Ay, x^*) \mid A^* x^* = y^* \} \right\} \\
= \sup_{(x^*,y^*,y)} \{ \langle y, v^* \rangle + \langle y^*, v \rangle - h (Ay, x^*) \mid A^* x^* = y^* \} \\
= \sup_{(x^*,y^*,y)} \{ \langle y, v^* \rangle + \langle y^*, v \rangle \\
- \inf \{ f (Ay, x_1^*) + g (Ay, x_2^*) \mid x_1^* + x_2^* = x^* \} \mid A^* x^* = y^* \}
\]

22
\[ = \sup_{(x_1^*, x_2^*, y^* \in \text{dom} V)} \left\{ \langle y, v^* \rangle + \langle y^*, v \rangle - f(Ay, x_1^*) - g(Ay, x_2^*) \mid A^* x_1^* + A^* x_2^* - y^* = 0 \right\} \]

\[ = -\max_{y \in Y} \inf_{(x_1, x_2, y, y^*)} \left\{ f(Ay, x_1^*) + g(Ay, x_2^*) - \langle y, v^* \rangle - \langle y^*, v \rangle + \langle y, A^* x_1^* + A^* x_2^* - y^* \rangle \right\} \]

where we have used the Lagrange multiplier rule [7, Cor. 4.4.4] to absorb the constraint \( A^* x_1^* + A^* x_2^* - y^* = 0 \) into the minimization. This is valid if

\[ V(x_1^*, x_2^*, y, y^*) := f(Ay, x_1^*) + g(Ay, x_2^*) - \langle y, v^* \rangle - \langle y^*, v \rangle \]

is lower semicontinuous and under the condition that \( 0 \in \text{core}(A(\text{dom} V)) \), where \( A(x_1^*, x_2^*, y, y^*) := A^* x_1^* + A^* x_2^* - y^* \in Y^* \). This condition is verified when \( \text{dom} V \neq \emptyset \), which is implied by (33) because

\[ A^{-1}(\text{Pr}_X(\text{dom } f)) \cap A^{-1}(\text{Pr}_X(\text{dom } g)) = \text{Pr}_Y(\text{dom } F) \cap \text{Pr}_Y(\text{dom } G) \neq \emptyset. \]

Continuing

\[ H^*(v^*, v) = -\max_{y \in Y} \inf_{(x_1^*, x_2^*, y, y^*)} \left\{ f(Ay, x_1^*) + \langle Ay, x_1^* \rangle + g(Ay, x_2^*) + \langle Ay, x_2^* \rangle - \langle y, v^* \rangle - \langle y^*, v + y \rangle \right\} \]

\[ = -\max_{y \in Y} \inf_{(x^*, y, y^*)} \left\{ \inf_{x_1^* + x_2^* = x^*} \left\{ f(Ay, x_1^*) + g(Ay, x_2^*) - \langle -Ay, x^* \rangle - \langle y, v^* \rangle - \langle y^*, v + y \rangle \right\} \right\} \]

\[ = \min_{y \in V} \sup_{y^*} \left\{ \sup_{(x^*, y)} \left\{ \langle -Ay, x^* \rangle + \langle y, v^* \rangle - \inf_{x_1^* + x_2^* = x^*} \left\{ f(Ay, x_1^*) + g(Ay, x_2^*) \right\} \right\} + \langle y^*, v + y \rangle \right\}. \]

Let \( F(y, x^*) := f(Ay, x^*) \) and \( G(y, x^*) := g(Ay, x^*) \). Now use [26, Thm 4.1] to evaluate

\[ \sup_{(x^*, y)} \left\{ \langle -Ay, x^* \rangle + \langle y, v^* \rangle - \inf_{x_1^* + x_2^* = x^*} \left\{ F(y, x_1^*) + G(y, x_2^*) \right\} \right\} \]

\[ = \min_{y_1^* + y_2^* = v} \left\{ F^*(y_1^*, -Ay) + G^*(y_2^*, -Ay) \right\} = (F^* \square_i G^*) (v^*, -Ay) \]

under the assumption that

\[ 0 \in \text{sqri} (\text{Pr}_Y (\text{dom } F) - \text{Pr}_Y (\text{dom } G)). \]

Now \( y \in \text{Pr}_Y (\text{dom } F) \) iff there exists \( x^{**} \) such that \( (Ay, x^{**}) \in \text{dom } f \) iff \( y \in A^{-1} (\text{Pr}_X (\text{dom } f) \) and so (34) is implied by

\[ 0 \in \text{sqri} [A^{-1} (\text{Pr}_X (\text{dom } f) - A^{-1} (\text{Pr}_X (\text{dom } g)]. \]
From (33) we have the existence of \( z = Ax \in \text{Range} A \cap \text{Pr}_X(\text{dom} g) \) such that \( X = \text{cone}(\text{Pr}_X(\text{dom} f) - z) \). Consequently we have a close subspace \( Z \) with

\[
Z := A^{-1}(X) = \text{cone} A^{-1}(\text{Pr}_X(\text{dom} f) - z) = \text{cone} [A^{-1}(\text{Pr}_X(\text{dom} f)) - x] \\
\subseteq \text{cone} [A^{-1} \text{Pr}_X(\text{dom} f) - A^{-1} \text{Pr}_X(\text{dom} g)] \subseteq \text{cone} A^{-1}(X) = Z,
\]

verifying (34). Finally

\[
H^* (v^*, v) = \min_{y^*} \sup_{\tilde{y}} \{(F^* \square_1 G^*) (v^*, -A\tilde{y}) + \langle y^*, v + \tilde{y} \rangle \} \\
= (F^* \square_1 G^*) (v^*, Av),
\]

due the attainment of the minimum at \( \tilde{y} = -v \).

To invoke the result [16, Prop. 1] or of [2] that ensure maximality of \( Q_A \) (on the reflexive space \( Y \)) we need to investigate the conjugate of \( H \). The next theorem outlines conditions that imply a sum theorem.

**Theorem 24** Suppose \( T \) and \( M \) are maximal monotone operators on a Banach space \( X \) and that \( A : Y \to X \) is an embedding of a finite-dimensional subspace \( \bar{Y} \subseteq X \) into \( X \). In addition suppose the following condition holds:

\[
\emptyset \neq \bar{Y} \cap \text{int} \text{ dom} T \cap \text{co dom} M. \tag{35}
\]

Let \( h := (f \square_2 g) \) and \( H (y, y^*) := \inf_{x^*} \{h (Ay, x^*) \mid A^* x^* = y^* \} \) with \( f = F_T, g = F_M \). Then

\[
H^* (y, y^*) \geq \langle y, y^* \rangle \quad \text{for all} \quad (y, y^*) \in Y \times Y^*. \tag{36}
\]

Consequently \( Q_A \) is maximal on \( Y \) and \( \overline{\text{dom} Q_A} \) is convex.

**Proof.** By Theorem 10 we know that \( T_A := A^* \circ T \circ A \) is maximal since \( \text{core} (\text{Pr}_X \text{ dom} F_T) \cap \text{Range} A \neq \emptyset \). Appealing to Corollary 16 we have for all \( (y^*, y) \in Y^* \times Y \) that \( F^* (y^*, Ay) \geq F_{A^* \circ T \circ A} (y^*, y) \geq \langle y, y^* \rangle \). Note that we do not know a priori that \( M_A := A^* \circ M \circ A \) is maximal, when \( A \) is the embedding of the finite-dimensional subspace \( \bar{Y} \) into \( X \). As \( Y \) is finite dimensional we know that \( Y \) is complemented by \( Z \), a close subspace, and so there exist continuous projections onto \( \bar{Y} \) and \( Z \) which we denote by \( P_Y : X \to \bar{Y} \) and \( P_Z : X \to Z \), respectively. In the following we will extend each \( y^* \in Y^* \) to \( \hat{y}^* \in X^* \) by defining \( \hat{y}^* (x) = y^* (x) \) if \( x \in Y \) and \( \hat{y}^* (x) = 0 \) if \( x \in Z \) and then extending via linearity to the whole space. Now use Proposition 23 to obtain \( y^*_1, y^*_2 \in Y \) such that \( y^* = y^*_1 + y^*_2 \)

\[
H^* (y, y^*) = (F^* \square_1 G^*) (y^*, Ay) \\
= F^* (y^*_1, Ay) + G^* (y^*_2, Ay)
\]

where \( F (y, x^*) = F_T (Ay, x^*) \) and \( G (y, x^*) = F_M (Ay, x^*) \).

24
Thus within \( \Pr H \) where the latter infimal convolution is exact and so \( F \) under the qualification assumption because the exactness of the convolution from (37) we have
\[
\mathcal{H}^* (y^*, x^{**}) := \left( \hat{\mathcal{F}} \square_\gamma \hat{G} \right) (y^*, x^{**}) : X^* \times X^{**} \to \mathbb{R}.
\]
where the latter infimal convolution is exact and so \( \mathcal{H}^* \) is proper convex. Now \( \Pr Y \) dom \( F = \Pr \gamma (\operatorname{dom} \mathcal{F}_T) \cap \hat{Y} \) along with \( \operatorname{int} \gamma \operatorname{Pr} Y \) dom \( \mathcal{F}_T \neq \emptyset \) due to (35). Thus within \( Y \) we have
\[
Y = \operatorname{cone}_Y \left( \Pr \gamma \operatorname{dom} \hat{F} - \Pr \gamma \operatorname{dom} \hat{G} \right)
\]
verifying (38). Restricting to \( \hat{X} \subseteq X^{**} \) we now consider the mapping
\[
(y^*, x) \mapsto \mathcal{H}^* (y^*, x) := \mathcal{H}^* (y^*, \hat{x}) .
\]
Next we note that
\[
\langle y^*, z \rangle = \langle \hat{y}^*, P_Y (z) \rangle = \langle y^*, P_Y (z) \rangle
\]
because \( z = P_Y (z) + P_Z (z) \) and hence
\[
\hat{F}^*(\hat{y}^*, x) = \sup_{(z,z^*) \in X \times X^*} \{ \langle \hat{y}^*, z \rangle + \langle x, z^* \rangle - \hat{F} (z, z^*) \} \\
= \sup_{(z,z^*) \in X \times X^*} \{ \langle \hat{y}^*, z \rangle + \langle x, z^* \rangle - \mathcal{F}_T (P_Y (z), z^*) \} \\
= \sup_{(P_Y (z), z^*) \in Y \times X^*} \{ \langle \hat{y}^*, P_Y (z) \rangle + \langle x, z^* \rangle - \mathcal{F}_T (P_Y (z), z^*) \} \\
= \sup_{(z,z^*) \in Y \times X^*} \{ \langle y^*, z \rangle + \langle x, z^* \rangle - \mathcal{F}_T (Ay, z^*) \} = \mathcal{F}^* (y^*, x) .
\]
From (37) we have \( H^* (y, y^*) = \mathcal{H}^* (\hat{y}^*, Ay) \) for all \( (y, y^*) \in Y \times X^* \). Using the exactness of the convolution
\[
\mathcal{H}^* (y^*, x) = F^* (y^*_1, x) + G^* (y^*_2, x) \\
= (F^* (y^*_1, x) - \langle y^*_1, x \rangle) + \langle y^*, x \rangle - \langle y^*_2, x \rangle + G^* (y^*_2, x) \\
\geq \langle y^*, x \rangle + (F^* (y^*_1, x) - \langle y^*_1, x \rangle) - \sup_{y^*_2 \in X^*} \{ \langle y^*_2, x \rangle - G^* (y^*_2, x) \} \\
= \langle \hat{y}^*, x \rangle + (F^* (y^*_1, x) - \langle \hat{y}^*_1, x \rangle) - \sup_{y^*_2 \in X^*} \{ \langle y^*_2, P_Y (x) \rangle - G^* (y^*_2, x) \} \\
= \langle \hat{y}^*, x \rangle + (F^* (y^*_1, x) - \langle \hat{y}^*_1, x \rangle) - (G^* (\cdot, x))^* (P_Y (x))
\]
where we denote
\[
(G^* (\cdot, x))^* (P_Y (x)) = \sup_{y^* \in X^*} \{ \langle y^*, P_Y (x) \rangle - G^* (y^*, x) \} .
\]
Now use Lemma 13 with $f = F_M$ and $p = P_M$ with $h(u, v) := -(P_M(u, \cdot))^*(v)$ (and a bounded linear operator $P_Y : X \to X$) to obtain via (19) that

$$G^*(y^*, x) = \sup_{z \in X} \{(z, y^*) + (h(\cdot, P_Y(z)))^{**}(x)\} = \sup_{z \in X} \{(z, y^*) - k(z, x)\}$$

where $k(z, x) := -(h(\cdot, P_Y(z)))^{**}(x)$. Consequently

$$(G^*(\cdot, x))^{**}(P_Y(x)) = (k(\cdot, x))^{**}(P_Y(x)) \leq k(P_Y(x), x) = -(h(\cdot, P_Y(x)))^{**}(x),$$

(40)

where $(h(\cdot, P_Y(x)))^{**}(x)$ is the lower semi-continuous hull of the function

$$u \mapsto h(u, P_Y(x)) = -(P_M(u, \cdot))^*(P_Y(x)) = \inf_{w^*} \{P_M(u, w^*) - \langle w^*, P_Y(x) \rangle\}. \quad (41)$$

Consequently

$$\mathcal{H}^*(y^*, x) \geq \langle \hat{y}^*, x \rangle + (F^*(\hat{y}_1^*, x) - \langle \hat{y}_1^*, x \rangle) - (k(\cdot, x))^{**}(P_Y(x)) \geq \langle \hat{y}^*, x \rangle + (F^*(\hat{y}_1^*, x) - \langle \hat{y}_1^*, x \rangle) + (h(\cdot, P_Y(x)))^{**}(x)$$

(42)\hspace{1cm} (43)

By Lemma 14 if $u \notin \overline{Pr_X \text{dom} P_M}$ then $P_M(u, w^*) = +\infty$ for all $w^*$ and hence $h(u, P_Y(x)) = +\infty$. In particular $(h(\cdot, P_Y(x)))^{**}(x) = +\infty$ for all $x \notin \overline{\text{clo} \text{ dom} M}$. Now consider $x, w \in \overline{Pr_X \text{dom} P_M}$ and on writing out the closure operations in detail we have (with limits taken with respect to the strong topology)

$$(k(\cdot, x))^{**}(w) = \lim_{z \to w} \inf k(z, x) = \lim_{z \to w} \inf \left( -(h(\cdot, P_Y(z)))^{**}(x) \right) = \lim_{z \to w} \left( \lim_{u \to x} \left( \inf_{\gamma \to \gamma_u} \left( \lim_{u \to x} \sup P_M(u, \cdot)^*(z_\gamma) \right) \right) \right) = \lim_{z \to w} \left( \inf_{\gamma \to \gamma_u} \left( \lim_{u \to x} \sup P_M(u, \cdot)^*(z_\gamma) \right) \right) \leq \inf_{\{z_u \to w\}} \lim_{u \to x} \sup P_M(u, \cdot)^*(z_u) = e-\lim_{u \to x} \sup P_M(u, \cdot)^*(w). \quad (44)$$

As $w \mapsto (k(\cdot, x))^{**}(w)$ is proper and lower semi-continuous convex we may take the conjugate twice and still preserve the inequality between the left and right hand sides, i.e.,

$$(k(\cdot, x))^{**}(w) \leq \left[ e-\lim_{u \to x} \sup P_M(u, \cdot)^* \right]^{**}(w). \quad (45)$$
Now we argue that $e\text{-lim sup}_{u \to x} (P_M (u, \cdot))^*$ never attains $-\infty$ (it is certainly lower semi–continuous and convex). If $e\text{-lim sup}_{u \to x} (P_M (u, \cdot))^*$ attains $-\infty$ then the right hand side of (45) would identically equal to $-\infty$ forcing $(k (\cdot, x))^{**} (\cdot) \equiv -\infty$. But then using (42) we have $H^* (y^*, \cdot) \equiv +\infty$, a contradiction to the properness of $H^*$. Note also that $P_M (u, w^*) = e_M^* (u, w^*)$ with the conjugate taken with the respect to the paired spaces $X \times X^*$ and $X^* \times X$ with $X$ endowed with the strong topology and $X^*$ endowed with the bounded weak* topology. Thus $(u, w^*) \mapsto P_M (u, w^*)$ is jointly $s \times bw^*$ lower semi–continuous and convex.

Thus

$$\left( bw^*-e\text{-lim inf}_{u \to x} P_M (u, \cdot) \right) (x^*) = \lim_{u^* \to x^*} \inf_{w^*} \left[ P_M (u, w^*) - P_M (u, x^*) \right]$$

with $x \mapsto P_M (x, x^*)$ a closed, convex function. Thus we may invoke Theorem 40 (see the Appendix) using the parametrised families $f\nu (\cdot) := (P_M (u, \cdot))^* (\cdot)$ and $f^\nu (\cdot) := P_M (u, \cdot)$ to deduce that

$$\left[ e\text{-lim sup}_{u \to x} (P_M (u, \cdot)) \right]^{**} (w) \leq \left[ bw^*-e\text{-lim inf}_{u \to x} P_M (u, \cdot) \right]^* (w). \quad (46)$$

Combining this with (42), (44) and (46) we obtain

$$H^* (y^*, x) \geq \langle \hat{y}^*, x \rangle + (F^* (y^*_1, x) - \langle \hat{y}^*_1, x \rangle) - (P_M (x, \cdot))^* (P_Y (x)). \quad (47)$$

When $x \in \hat{Y} \cap (Pr_X \text{ dom } P_M)$ we have $x = P_Y (x)$ and so using $\langle w^*, x \rangle - P_M (x, w^*) \leq 0$ for all $w^*$ we have

$$(P_M (x, \cdot))^* (x) = \sup_{w^*} \{ \langle w^*, x \rangle - P_M (x, w^*) \} \leq 0. \quad (48)$$

Combining (48) with (47) we have for $x \in \hat{Y} \cap (Pr_X \text{ dom } P_M)$

$$H^* (y^*, x) \geq \langle \hat{y}^*, x \rangle + (F^* (y^*_1, x) - \langle \hat{y}^*_1, x \rangle).$$

Now we use Proposition 23 to obtain $H^* (y, y^*) = H^* (y^*, Ay)$ for all $(y, y^*) \in Y \times Y^*$. Thus using Lemma 15, Corollary 16, Theorem 10 and init dom $T \neq \emptyset$ we have

$$H^* (y, y^*) = H^* (\hat{y}^*, Ay) \geq \langle \hat{y}^*, Ay \rangle + (F^* (y^*_1, Ay) - \langle \hat{y}^*_1, Ay \rangle)$$

$$\geq \langle A^* \hat{y}^*, y \rangle + (F_{A^*} (y^*_1, y) - \langle A^* \hat{y}^*_1, y \rangle)$$

$$= \langle y^*, y \rangle + (F_{A^*} (y^*_1, y) - \langle y^*_1, y \rangle) \geq \langle y^*, y \rangle$$

whenever $\hat{y} \in (Pr_X \text{ dom } P_M) \cap \hat{Y}$. When $x \notin Pr_X \text{ dom } P_M$ then $(P_M (x, \cdot))^* (P_Y (x)) = -\infty$ and via (47) we have $H^* (y^*, x) = +\infty$. Consequently $H^* (y, y^*) = H^* (y^*, Ay) \geq \langle y^*, y \rangle$ for all $(y, y^*) \in Y \times Y^*$.

The final assertion follows from the observation that all maximal monotone operators on reflexive spaces have semiconvex domains and that maximality of $Q_A$ follows from the converse of (6) proved in [16].

The representative function $H$ will now be used to define a maximal monotone set $Q_A$ which we must compare with $(T + M)_A$. Recall (32):
Proposition 25 Suppose $T$ and $M$ are maximal monotone operators on a Banach space $X$ and that $A : Y \to X$ is an embedding of a finite-dimensional subspace $Y \subseteq X$ into $X$. Define a monotone mapping $Q_A$ by placing $h := (\square_2 g)$, $H (y, y^*) := \inf_x \{ h (Ay, x^*) | A^* x^* = y^* \}$ and defining the graph of $Q_A$ as

$$Q_A := \{ (y, y^*) | H (y, y^*) = \langle y, y^* \rangle \}.$$  

Suppose

$$\emptyset \neq (\Pr_X \dom P_M) \cap \interior (\Pr_X \dom P_T) \cap \hat{Y}.$$  

Then for all $y \in \interior (\dom T) \cap \hat{Y}$ and any $y^* \in Y^*$ we have

$$Q_A (y) \subseteq (T + M)_A (y).$$  

(50)

Proof. Taking $x_0 \in (\Pr_X \dom P_M) \cap \interior (\Pr_X \dom P_T)$ and translating the origin to $x_0$ we may assume with out loss of generality that $0 \in (\Pr_X \dom P_M) \cap \interior (\Pr_X \dom P_T)$. Note that since the qualification assumption implies

$$0 \in \core [(\Pr_X \dom P_T) - (\Pr_X \dom P_M)],$$

we have $h := f \square_2 g$ is exact. Then $H (y, y^*) = \langle y, y^* \rangle$ exactly when

$$0 = \inf_{v \ast} \{ f (Ay, v_2) - \langle y, A^* v_2 \rangle + g(Ay, x^* - v_2) - \langle y, A^* (x^* - v_2) \rangle | A^* x^* = y^* \}$$

$$= \inf_{(v_1, v_2)} \{ (f (Ay, v_1) - \langle y, v_1 \rangle) + (g(Ay, v_2) - \langle Ay, v^*_2 \rangle) | A^* (v_1 + v_2) = y^* \},$$

(51)

where both terms inside the infimum are positive. Now $Ay \in \Pr_X (\square T) \cap \Pr_X (\square M)$, otherwise $f (Ay, v^*_1) + g(Ay, v^*_2) > \langle Ay, v^*_1 + v^*_2 \rangle$ irrespective of the choice of $v^*_1, v^*_2$. Consequently the domain of $Q_A$ is contained in $\Pr_X (\square T) \cap \Pr_X (\square M) \cap \hat{Y}$ (a convex set in a finite-dimensional subspace). Denote the standard $\varepsilon$-enlargement, for $\varepsilon > 0$, by

$$M_\varepsilon := \{ (x, x^*) | g(x, x^*) - \langle x, x^* \rangle \leq \varepsilon \}$$

with $T_\varepsilon$ defined similarly. We now observe that that by Lemma 2.2 part c) of [24] and that fact that $f = (f^*)^\ast$ (the conjugates taken with respect to the paired spaces $\sigma_{w^*} \times (X^* \times X)$ and $\sigma_{w} \times (X \times X^*)$) that for all $s_0 \in \Pr_X \dom P_T \subseteq \Pr_X \dom f^\ast$ we have the existence of $K > 0, \eta \in (0, 1]$ such that $\| s - s_0 \| \leq \eta$ and $(y, y^*) \in X \times X^*$ we have $f (y, y^*) + K \| y - s \| - (s, y^*) \geq \eta (\| y^* \| - K)$. Consequently for all $s^* \in T_\varepsilon (s)$ we have (on letting $y = s$ and $y^* = s^*$)

$$\varepsilon \geq \eta (\| s^* \| - K) \implies \diam T_\varepsilon (s) \leq K + \frac{\varepsilon}{\eta} \text{ for all } \| s - s_0 \| \leq \eta.$$  

Hence $T_\varepsilon$ is locally bounded within $\interior \Pr_X \dom P_T$.

Take a minimizing net $\{ (v_1^\ast, v_2^\ast) \}$ for (51) and note that both terms in (51) must tend to zero. Equation (51) implies for all $v_2^\ast \in M_\varepsilon \circ A (y)$ such that

$$A^* (v_1^\ast + v_2^\ast) = y^*$$

we have

$$f (Ay, v_1^\ast) - \langle Ay, v_1^\ast \rangle - \beta 0.$$  

(52)
Consequently there is a $\beta_\varepsilon$ such that for $\beta \geq \beta_\varepsilon$ we have $v_1^{\ast \beta} \in T_\varepsilon \circ A (y)$ and $v_1^{\ast \beta} \in T_\varepsilon \circ A (y)$. Let $A^* v_2^{\ast \beta} \leq y_2^{\ast \beta}$ then for all $y_2^{\ast \beta} \in (A^* \circ M_\varepsilon \circ A) (y) = (M_\varepsilon) A (y)$ such that $A^* v_1^{\ast \beta} = y^* - y_2^{\ast \beta} \in (A^* \circ T_\varepsilon \circ A) (y) := (T_\varepsilon) A (y)$ we have (52). Now (52) implies

$$\sup_{(x, x^*) \in T} \langle Ay - x, x^* - v_1^{\ast \beta} \rangle \rightarrow 0.$$ 

Hence for all $(x, x^*) \in T$ and for $\beta \geq \beta_\varepsilon \vee \beta_\varepsilon$ such that $\sup_{(x, x^*)} \in T (Ay - x, x^* - v_1^{\ast \beta}) \leq \varepsilon$ it follows that

$$\langle v_1^{\ast \beta}, x \rangle \leq \langle x, x^* \rangle - (Ay, x^*) + \langle Ay, v_1^{\ast \beta} \rangle + \varepsilon$$

$$= \langle x, x^* \rangle - (Ay, x^*) + (y, A^* v_1^{\ast \beta}) + \varepsilon$$

Taking convex combinations and limits we have for all $x \in \Pr_X (\text{dom } \mathcal{P}_T), \beta \geq \beta_\varepsilon \vee \beta_\varepsilon$ and $(x, x^*) \in \text{dom } \mathcal{P}_T$ that

$$\langle v_1^{\ast \beta}, x \rangle \leq \mathcal{P}_T (x, x^*) - (Ay, x^*) + (y, y^* - y_2^{\ast \beta}) + \varepsilon$$

$$\leq \mathcal{P}_T (x, x^*) - (Ay, x^*) + S((T_\varepsilon) A (y), y) + \varepsilon := K (x, x^*, A, y).$$

As $0 \in \text{int \ Pr}_X \text{ dom } \mathcal{P}_T \neq \emptyset$ we have (53) implying a bound for all $x \in X = \text{cone} \text{ Pr}_X \text{ dom } \mathcal{P}_T$ where the right hand side bound is independent $\{v_1^{\ast \beta}\}_\beta$.

Thus, by the uniform boundedness principle we have $\{v_1^{\ast \beta}\}_\beta$ bounded whenever $(T_\varepsilon) A (y)$ is bounded. Using [24, Thm 3.6] ensures that

$$\text{int \ dom } T = \text{int \ Pr}_X \text{ dom } \mathcal{F}_T \supseteq \text{int \ Pr}_X \text{ dom } \mathcal{P}_T \supseteq \text{int \ dom } T.$$ 

It follows that $\text{int \ Pr}_X \text{ dom } \mathcal{P}_T = \text{int \ dom } T$ and within this set $T$ is locally bounded, and so $(T_\varepsilon) A (y)$ is bounded for $Ay \in \text{int \ dom } T$.

Consequently, for these $y$ we have $\{v_1^{\ast \beta}\}$ bounded and we may take a subnet and obtain $v_1^{\ast \beta} \rightarrow_w^* v_1^{\ast \beta} \in T \circ A (y)$. It follows that $A^* v_1^{\ast \beta} \rightarrow^* y_1^{\ast \beta} \in (A^* \circ T_\varepsilon \circ A) (y) := (T_\varepsilon) A (y)$ and as $A^* v_1^{\ast \beta} = y^* - y_2^{\ast \beta}$ it also follows that $\{y^* - y_2^{\ast \beta}\}_\beta$ converges. Let $w^* \text{- } \lim_{\beta} y_2^{\ast \beta} = y_2^*$ and so $y^* - y_2^* = y_1^*$.

Thus, there exists $y_2^{\ast \beta} \in (M_\varepsilon) A (y)$ with $w^* \text{- } \lim_{\beta} y_2^{\ast \beta} = y_2^*$ and by the weak* closedness of $(M_\varepsilon) A (y)$ we have $y_2^* \in (M_\varepsilon) A (y)$ where $\varepsilon > 0$ is arbitrary. Consequently,

$$y^* - y_1^* \in \cap_{\varepsilon>0} (M_\varepsilon) A (y) = A^* \circ (\cap_{\varepsilon>0} M_\varepsilon) \circ A (y)$$

$$= (A^* \circ M \circ A) (y) = M_A (y).$$

Similarly $v_1^{\ast \beta} \rightarrow_w^* v_1 \in T \circ A (y)$ and $A^* v_1^{\ast \beta} = y^* - y_2^{\ast \beta}$ we have $A^* v_1 = y_1^* = y^* - y_2^*$ with $y_1^* \in T_A (y)$. Thus

$$y^* \in T_A (y) + M_A (y) = (T + M)_A (y).$$

29
We have shown that for all
\[ \hat{y} = Ay \in \text{int} (\text{Pr}_X \text{dom} \mathcal{P}_T) \cap \hat{Y} = \text{int} (\text{dom} T) \cap \hat{Y} \]
it follows that
\[ H (y, y^*) = \langle y, y^* \rangle \Rightarrow y^* \in (T + M)_A (y). \]
Thus, (50) holds and we are done. \( \blacksquare \)

The containment of the monotone relations \( Q_A \) and \( (T + M)_A \) may only fail on the boundary of \( \text{dom} T \). By definition it is clear that
\[ \text{dom} (T + M)_A = \text{dom} M \cap \hat{Y} \cap \text{dom} T. \]
We need to extend this inclusion to all of \( \text{dom} Q_A \) and to do so we need to characterise its domain. This will occupy our attention for next few results under the following assumption.

\( \text{(A)} \): Assume \( T \) and \( M \) are maximal monotone operators on a Banach space \( X \) and that \( A : Y \to X \) is an embedding of a finite-dimensional subspace \( \hat{Y} \subseteq X \) into \( X \). Define a monotone mapping \( Q_A \) via the representative function \( H \) as in Proposition 25 with \( f (x, x^*) = \mathcal{F}_T (x, x^*) \) and \( g (x, x^*) = \mathcal{F}_M (x, x^*) \) and suppose \( \text{dom} M \cap \hat{Y} \cap \text{int} \text{dom} T \neq \emptyset \) and so
\[ \text{dom} M \cap \text{int} \text{dom} T \neq \emptyset \] (55)

**Remark 26** Note that assumption (A) implies the assumption (35) of Theorem 24 and also (49) of Proposition 25.

**Corollary 27** Suppose we have assumption (A) holding then
\[ \text{dom} M \cap \hat{Y} \cap \text{dom} T \subseteq \text{dom} Q_A \quad \text{and} \]
\[ (\text{int} \text{dom} T) \cap \overline{\text{dom} Q_A} \subseteq \overline{\text{dom} M \cap \hat{Y} \cap \text{int} \text{dom} T}. \]

**Proof.** From Proposition 25 we have
\[ (\text{int} \text{dom} T) \cap \text{dom} Q_A \subseteq \text{dom} (T + M)_A = \text{dom} M \cap \hat{Y} \cap \text{dom} T. \] (57)
Let \( Ay \in \text{dom} T \cap \text{dom} M \cap \hat{Y} \). First observe that even when the support of some \( x^* \in (T + M) \circ A (y) \) does not intersect \( Y \) we still have \( A^* x^* = x^* |_Y = 0 \) and so for \( y^* = 0 \) (using that fact that \( H \) is a representative function (30)),
\[ 0 = \langle y, y^* \rangle \leq H (y, y^*) = \inf \{ h (Ay, \hat{x}^*) \mid A^* \hat{x}^* = 0 \} \]
\[ \leq h (Ay, x^*) = \langle Ay, x^* \rangle = \langle y, A^* x^* \rangle = 0 = \langle y, y^* \rangle \]
giving \( y \in \text{dom} Q_A \) with \( Q_A (y) = \{ 0 \} \).
Otherwise, for any \( x^* \in (T + M) \circ A (y) \), with \( y^* = x^*|_Y \) we have
\[
\langle y, y^* \rangle \leq H (y, y^*) = \inf_{z^*} \{ h (Ay, z^*) | A^* z^* = y^* \}
\]
\[
\leq h (Ay, x^*) = \langle Ay, x^* \rangle = \langle y, A^* x^* \rangle = \langle y, y^* \rangle
\]
implying \( y^* \in Q_A (y) \) i.e. \( y \in \text{dom} Q_A \) and
\[
Q_A (y) \supseteq \{ y^* = x^*|_Y | x^* \in (T + M) \circ A (y) \} = (T + M)_A (y).
\]
Consequently, \( Q_A \supseteq \text{dom} M \cap \hat{Y} \cap \text{dom} T \). Combining this observation with (57) we have the first two inclusions (56). For the last inclusion we take \( x \in (\text{int dom} \ T) \cap \text{dom} Q_A \) and so there exists \( x_n \to x \) with \( x_n \in \text{dom} Q_A \). Thus eventually \( x_n \in \text{int dom} T \) and so \( x_n \in (\text{int dom} \ T) \cap \text{dom} Q_A \subseteq \text{dom} M \cap \hat{Y} \cap \text{int dom} T \) by (57). Taking limits gives the last inclusion of (56).

We need the exact form of \( \text{dom} Q_A \) and \( \text{dom} H \) under the following standing assumption (A):

**Proposition 28** Suppose we have assumption (A) holding then
\[
\text{Pr}_X \text{dom} H \supseteq (\text{co dom} M) \cap \hat{Y} \cap \text{dom} T = (\text{Pr}_X \text{co} M) \cap \hat{Y} \cap \text{dom} T,
\]
and so when \( Q_A \) is maximal
\[
\text{Pr}_X \text{dom} H = \text{dom} Q_A \supseteq (\text{co dom} M) \cap \hat{Y} \cap \text{dom} T.
\]

**Proof.** First observe that \( \text{dom} M \cap \hat{Y} \subseteq \text{Pr}_X \text{dom} F_M \) and so \( \text{co} (\text{dom} M \cap \hat{Y}) \subseteq \text{Pr}_X \text{dom} F_M \) (because \( F_M \) is a convex function) implying by (31)
\[
\text{Pr}_X \text{dom} H \supseteq \text{co} (\text{dom} M \cap \hat{Y}) \cap \text{dom} T
\]
\[
= \text{co} (\text{dom} M \cap \hat{Y}) \cap \text{dom} T
\]
where the last equality of (58) follows from \( \text{co} (\text{dom} M \cap \hat{Y}) \cap \text{int dom} T \neq \emptyset \) and standard convex analysis results.

When \( Q_A \) is maximal on \( Y \) (a finite dimensional (reflexive) space) by the semi–convexity of \( \text{dom} Q_A \) we have \( \text{dom} Q_A \) convex and
\[
\text{dom} Q_A = \text{Pr}_X \text{dom} H \supseteq \text{co} (\text{dom} M \cap \hat{Y}) \cap \text{dom} T.
\]

The domain of \( Q_A \) could differ from that of \( (T + M)_A \), an issue that needs to be resolved at least for the case when \( f (x, x^*) = F_T (x, x^*) \) and \( g (x, x^*) = F_M (x, x^*) \).
Corollary 29 Suppose we have assumption (A) holding. Then

\[ \text{dom } Q_A \subseteq \text{Pr}_X \text{ dom } H \subseteq \overline{\text{dom } T \cap \hat{Y}} = \text{co} \left( \overline{\text{dom } T \cap \hat{Y}} \right). \]

Consequently

\[
\begin{align*}
\text{dom } Q_A &= \overline{\text{dom } (T + M)_A} \\
&= \text{dom } M \cap \hat{Y} \cap \text{int dom } T = \text{co} \left( \text{dom } M \cap \hat{Y} \cap \text{int dom } T \right) \\
&= \left( \text{co dom } M \right) \cap \hat{Y} \cap \text{dom } T. \tag{60}
\end{align*}
\]

Proof. Using the semi-convexity property for \( T \) (since \( T \) is maximal monotone and \( \text{int dom } T \neq \emptyset \), see [24, Thm 3.6, Thm 3.8]) we have \( \overline{\text{co } \text{Pr}_X T} = \overline{\text{Pr}_X T} \) and so using Lemma 12

\[
\begin{align*}
A^{-1} \left( \overline{\text{Pr}_X T} \right) &\subseteq \text{co } A^{-1} \left( \overline{\text{Pr}_X T} \right) \subseteq A^{-1} \left( \overline{\text{Pr}_X \text{co } T} \right) = A^{-1} \left( \text{co } \text{Pr}_X T \right) \\
&= A^{-1} \left( \overline{\text{Pr}_X T} \right) \subseteq \text{co } \left( A^{-1} \left( \overline{\text{Pr}_X T} \right) \right)
\end{align*}
\]

where the second inclusion follows from the observation that the set \( \text{co } A^{-1} \left( \overline{\text{Pr}_X T} \right) \) is the smallest convex set containing \( A^{-1} \left( \overline{\text{Pr}_X T} \right) \) and \( A^{-1} \left( \overline{\text{Pr}_X \text{co } T} \right) \) is a convex set containing \( A^{-1} \left( \overline{\text{Pr}_X T} \right) \). Thus we have the equalities

\[
\begin{align*}
\text{co } A^{-1} \left( \overline{\text{Pr}_X T} \right) &= \text{co } \left( \overline{\text{Pr}_X T \cap \hat{Y}} \right) = \text{co } \left( \overline{\text{dom } T \cap \hat{Y}} \right) \\
&= A^{-1} \left( \overline{\text{Pr}_X T} \right) = A^{-1} \left( \text{dom } T \right) = \overline{\text{dom } T \cap \hat{Y}}.
\end{align*}
\]

Using [24, Thm 3.8] which states that

\[ \text{Pr}_X \text{ dom } F_T = \text{int } (\text{Pr}_X \text{ dom } F_T) = \text{int dom } T = \overline{\text{dom } T} \tag{62} \]

and combining this with (31) we have \( \text{Pr}_X \text{ dom } H \subseteq \overline{\text{dom } T \cap \hat{Y}} \). In particular

\[ \overline{\text{dom } Q_A} = \text{dom } Q_A \cap \text{dom } T. \]

Via maximality \( \overline{\text{dom } Q_A} \) is convex, using the conclusion of Theorem 24 it follows that

\[
\begin{align*}
\overline{\text{dom } Q_A} &= \text{co } (\text{dom } Q_A) \supseteq \text{co } \left( \text{dom } M \cap \hat{Y} \cap \text{dom } T \right) \tag{63} \\
&= \text{co } \left( \text{dom } (T + M)_A \right) \supseteq \text{co } \left( \text{dom } M \cap \hat{Y} \cap \text{int dom } T \right).
\end{align*}
\]

By Theorem 2.13 of [6], the fact that \( \text{qri } A = \text{ri } A \) if \( A \) is finite dimensional, that fact that \( \text{dom } Q_A \cap \text{int dom } T \neq \emptyset \) and \( \text{ri } \text{dom } Q_A \neq \emptyset \) (because \( \text{dom } Q_A \) is a convex set in finite dimensions) we have

\[
\text{ri } \text{dom } Q_A = \text{ri } (\text{dom } Q_A \cap \text{dom } T) = \text{ri } \text{dom } Q_A \cap \text{int dom } T \\
\subseteq \text{dom } Q_A \cap \text{int dom } T \subseteq \text{dom } M \cap \hat{Y} \cap \text{int dom } T \subseteq \overline{\text{dom } (T + M)_A},
\]

32
where we have used the last inclusion of Corollary 27. Thus
\[ \text{dom} Q_A = \text{ri} \text{dom} Q_A \subseteq (\text{dom} M \cap \hat{Y} \cap \text{int dom} T). \]
Combining this with (56), (63) gives (60). As
\[ (\text{co dom} M) \cap \hat{Y} \cap \text{dom} T \supseteq \text{co} (\text{dom} M \cap \hat{Y} \cap \text{int dom} T) \]
then (59) gives the equality (61).

The following give a first step toward semi-convexity of \( \text{dom} M \) for a arbitrary maximal monotone mapping on a nonreflexive space.

**Corollary 30** Suppose we have assumption (A) holding. Then
\[ \text{co} (\text{dom} M \cap \hat{Y}) = \text{dom} M \cap \hat{Y} = (\text{co dom} M) \cap \hat{Y}. \] (64)

Consequently we have
\[ \text{dom} Q_A = \text{dom} M \cap \hat{Y} \cap \text{dom} T \]
\[ = \text{dom} M \cap \hat{Y} \cap \text{dom} T \cap \hat{Y} = (\text{co dom} M) \cap \hat{Y} \cap \text{dom} T. \] (65)

**Proof.** Using a maximal monotone operator \( T \) with \( \text{dom} T = X \), such as \( T(x) \equiv 0 \), we apply Corollary 29 to obtain (64). In particular \( \text{dom} M \cap \hat{Y} \) is convex. Now take some other monotone operator \( T \) which only has \( \text{int dom} T \neq \emptyset \). When \( \text{dom} M \cap \hat{Y} \cap \text{int dom} T \neq \emptyset \) we have
\[ \text{dom} M \cap \hat{Y} \cap \text{int dom} T = \text{dom} M \cap \hat{Y} \cap \text{dom} T \]
As \( \text{int dom} T \cap \hat{Y} \neq \emptyset \) we have \( \text{dom} T \cap \hat{Y} = \text{dom} T \cap \hat{Y} \) which will give the second equality in (65) after the following is observed. Take \( x \in \text{dom} M \cap \hat{Y} \cap \text{int dom} T \) then there exists \( x_n \rightarrow x \) with \( x_n \in \text{dom} M \cap \hat{Y} \cap \text{int dom} T \). For \( n \) large \( x_n \in \text{int dom} T \) and \( x_n \in \text{dom} M \cap \hat{Y} \) implies the existence of \( z_m \rightarrow x_n \) with \( z_m \in \text{dom} M \cap \hat{Y} \). A diagonalisation argument gives \( z_{m_k} \rightarrow x \) with \( z_{m_k} \in \text{dom} M \cap \hat{Y} \cap \text{int dom} T \). Thus \( x \in \text{dom} M \cap \hat{Y} \cap \text{int dom} T \) and hence
\[ \text{dom} M \cap \hat{Y} \cap \text{int dom} T = \text{dom} M \cap \hat{Y} \cap \text{int dom} T = \text{dom} Q_A. \]

These results culminate in a generalization of Rockafellar’s sum theorem [19].

**Theorem 31 (Sum Theorem)** Suppose \( T \) and \( M \) are maximal monotone operators on a Banach space \( X \). In addition suppose
\[ \emptyset \neq \text{dom} M \cap \text{int dom} T. \] (66)

Then \( T + M \) is a maximal monotone operator.
Proof. Our assumptions imply \( 0 \in \text{core} (\Pr_X \text{dom } \mathcal{P}_T - \Pr_X \text{dom } \mathcal{P}_M) \) and \( \emptyset \neq \text{core} (\text{dom } \mathcal{P}_T) \subseteq \text{core} (\Pr_X \text{dom } \mathcal{F}_T) \) (as \( \text{dom } \mathcal{P}_T \subseteq \text{dom } \mathcal{F}_T \)) and hence \( \text{core} (\Pr_X \text{dom } \mathcal{F}_T) \cap \Pr_X \text{dom } \mathcal{F}_M \neq \emptyset \). We apply Theorem 8 but first make a suitable translation. Take \( x \in \text{core} \Pr_X \text{dom } \mathcal{F}_T \cap \Pr_X \text{dom } \mathcal{F}_M \) and place \( \hat{T} = T \cdot (-x) \) and \( \hat{M} = M \cdot (-x) \). It is easily shown that

\[
\mathcal{F}_{\hat{T} \cdot (-z,z^*)} (y, y^*) = \mathcal{F}_T (y + z, y^* + z^*) - \langle (y + z, y^* + z^*) - (y, y^*) \rangle.
\]

Hence, on identifying the translations with the original operators we may assume

\[
0 \in \text{core} (\Pr_X \text{dom } \mathcal{F}_T) \quad \text{and that} \quad 0 \in (\Pr_X \text{dom } \mathcal{F}_T) \cap (\Pr_X \text{dom } \mathcal{F}_M).
\]

As \( T \) and \( M \) are maximal, \( f := \mathcal{F}_T \) and \( g := \mathcal{F}_M \) are both representative functions and since \( \hat{f}^* = \hat{f}^*_T \) (resp. \( \hat{g}^* = \hat{g}^*_M \)) we have \( \hat{f}^* \geq \langle \cdot \rangle \) (resp. \( \hat{g}^* \geq \langle \cdot \rangle \)) on \( X^* \times X^{**} \) (as observed in Theorem 3). Thus, by Theorem 21 with \( h := (f \square_2 g) \), we deduce that \( \hat{h}^* = \left( \hat{f}^* \square_1 \hat{g}^* \right) (x^*, x) \) is a representative functions with \( \langle h^* \rangle \) being a representative function for \( T + M \) and with \( \hat{h}^* \geq \langle \cdot \rangle \) on \( X^* \times X^{**} \).

In order to apply Theorem 8 (or rather Remark 9) we first show that for every finite-dimensional subspace \( Y \) of \( X \) such that \( Y \cap \text{core} (\Pr_X \text{dom } \mathcal{F}_T) \neq \emptyset \) we have

\[
(T + M)_A := A^* \circ (T + M) \circ A
\]

maximal where \( A : Y \rightarrow X \) is the embedding of \( Y \) into \( X \). As \( Y \) is reflexive it suffices to show that there exists a representative function \( H \) of \( (T + M)_A \) such that \( H \geq \langle \cdot \rangle \) and \( H^* \geq \langle \cdot \rangle \) on \( Y \times Y^* \) and as proved in Proposition 1 of [16]. To this end, we take \( h := (f \square_2 g) \) and \( H (y, y^*) := \inf_{x} \{ h (Ay, x^*) \mid A^* x^* = y^* \} \). By Theorem 21 since \( h \) is representative, \( H \) is representative, that is,

\[
H (y, y^*) \geq \inf_{x} \{ h (Ay, x^*) \mid A^* x^* = y^* \} \geq \langle y, y^* \rangle.
\]

By Proposition 24 we have \( H^* (y, y^*) \geq \langle y, y^* \rangle \) and hence \( H^* \) is a representative function. By [16, Prop. 1] and the reflexivity of \( Y \) we know that \( H \) is a representative function of a maximal monotone operator given by \( Q_A \) as defined in Proposition 25.

Clearly we can cover \( X \) with such finite-dimensional subspaces \( Y \)—a condition like \( Y \cap \text{core} (\Pr_X \text{dom } \mathcal{F}_T) \neq \emptyset \) is easily satisfied by taking any \( x \in \text{int} \text{dom } T \cap \text{dom } M \) and forming \( Y := \text{span} (Y' \cup \{ x \}) \) where \( Y' \) can be any finite-dimensional subspace. Now we have

\[
\text{dom } Q_A = \overline{(\text{co dom } M) \cap Y \cap \text{dom } T}. \quad (68)
\]

We now restrict attention to the subspace \( Y \). By Theorem 22 the function \( h \) is representative for \( T + M \). By Corollary 30 we have

\[
\text{dom } Q_A = \overline{(\text{co dom } M) \cap Y \cap \text{dom } T}. \quad (68)
\]

34
As \((\text{co dom } M \cap Y) \cap \text{int (dom } T) \neq \emptyset\) we may apply Theorem 2.13 of [6]—using \(\text{qri } C = \text{ri } C\) when \(C\) is finite-dimensional—to deduce that

\[
\text{ri dom } Q_A = \text{ri} \left( (\text{co dom } M) \cap \hat{Y} \cap \text{dom } T \right)
= \left( \text{ri} \left( (\text{co dom } M) \cap \hat{Y} \right) \cap \text{int dom } T \right) \subseteq \left[ \text{int (dom } T) \right] \cap \hat{Y}.
\]

By Proposition 25 we know \(Q_A (y) \subseteq (T + M)_A (y)\) for all \(y \in \left[ \text{int (dom } T) \right] \cap \hat{Y}\) and any \(y^* \in Y^*\). Thus the inclusion \(Q_A \subseteq (T + M)_A\) holds on \(\text{ri dom } Q\). Now apply Corollary 19 to deduce

\[
Q_A (y) = \overline{\text{co}} \left( \limsup_{y' \in \text{ri dom } Q_A} Q_A (y') + N_{\text{dom } Q_A} (y) \right).
\]

(69)

We now use the fact that \((T + M)_A\) has closed graph [5]. This follows from the observation that \(T + M = \overline{T + M}\) since its representative function \(f := (g \square_2 h)\) is \(s \times bw^*\) closed and

\[
T + M = \{ (x, x^*) : (g \square_2 h) (x, x^*) - \langle x, x^* \rangle \leq 0 \}.
\]

Consequently,

\[
(T + M)_A := (T + M) \cap (Y \times (X^*/Y^*))
\]

is a closed set in the finite-dimensional subspace \(Y \times (X^*/Y^*) \simeq Y \times Y^*\) (see [14, pp. 123]). Thus, \((T + M)_A\) has closed convex images. Using this observation and (50)

\[
\text{co lim sup}_{y' \to y} Q_A (y') \subseteq \text{lim sup}_{y' \to y} (T + M)_A (y')
\]

(70)

\[
\subseteq (T + M)_A (y).
\]

Let \(\hat{y} \in \text{bd } \overline{\text{dom } Q_A} = \text{bd } \left( \text{dom } T \cap \hat{Y} \cap \text{dom } M \cap \hat{Y} \right)\). As \(\text{dom } M \cap \text{int (dom } T) \cap \hat{Y} \neq \emptyset\) we have \(0 \in \text{core}_{\hat{Y}} \left( \text{dom } T \cap \hat{Y} - \text{dom } M \cap \hat{Y} \right)\) and so

\[
N_{\text{dom } Q_A} (y) = N_{\text{dom } M \cap Y \cap \text{dom } T \cap \hat{Y}} (y) = N_{(\text{dom } T) \cap \hat{Y}} (y) + N_{(\text{dom } M) \cap \hat{Y}} (y).
\]

(71)

By Corollary 19, Lemma 20, (71) and (70)

\[
Q_A (y) = \overline{\text{co}} \left( \limsup_{y' \to y} N (y') + N_{\text{dom } Q_A} (y) \right)
\]

\[
\subseteq T_A (y) + N_{(\text{dom } T) \cap \hat{Y}} (y) + M_A (y) + N_{(\text{dom } M) \cap \hat{Y}} (y)
\]

\[
\subseteq T_A (y) + (N_{(\text{dom } T)} A (y) + M_A (y) + (N_{(\text{dom } M)} A (y)
\]

\[
= T_A (y) + M_A (y) = (T + M)_A (y).
\]

35
Since $Q_A$ is maximal, equality ensues and proves the maximality of $(T + M)_A$.

By Theorem 8 we have $F_{M_*} : X \times X^* \to \mathbb{R}$ a representative function and $M_h = M_{h^*}$. Using Theorem 22 we have $h$ is a representative function of $T + M$. Thus, we have $M_{h^*} = T + M$ and so by Theorem 8, $F_{T+M}$ is a representative function. We will now argue as in Proposition 5 of [5]. Suppose $(x, x^*)$ is monotonically related to $T + M$ then $F_{T+M} (x, x^*) \leq \langle x, x^* \rangle$ and as $F_{T+M}$ is a representative function $F_{T+M} (x, x^*) = \langle x, x^* \rangle$. But by Proposition 2 of [5] we must have $P_{T+M} (x, x^*) = \langle x, x^* \rangle$ and so using Theorem 21

$$\langle x, x^* \rangle = P_{T+M} (x, x^*) \geq h^* (x^*, x) = (f^\star \Box_1 g^*) (x^*, x) \geq \langle x, x^* \rangle.$$ 

As $h^*$ is a representative of $T + M$ we have $x^* \in (T + M) (x)$, completing the proof.

By the same methods we may also prove the following composition result.

**Theorem 32 (Composition)** Suppose $X$ and $Y$ are Banach spaces, that $T$ is a maximal monotone operator on $Y$, and that $A : X \to Y$, is a bounded linear mapping. Then $T_A := A^* \circ T \circ A$ is maximal monotone on $X$ whenever $0 \in \text{range}(A) + \text{int dom} T$.

We refer to [32, Thm 6], [3, 4] for results when $Y$ is assumed reflexive. A special case is worth recording.

**Corollary 33 (Normal Cones.)** Suppose in an arbitrary Banach space that $T$ is maximal monotone and $C$ is closed and convex while either

$$C \cap \text{int } D(T) \neq \emptyset \quad \text{or} \quad D(T) \cap \text{int } C \neq \emptyset.$$ 

Then $T + N_C$ is maximal monotone.

We finish this section with a corollary extending one in [5] and which answers a quite long-standing open question. Recall that a maximal monotone mapping $T$ is maximal monotone locally [27] or of type (FPV), if for every open set $V$ in $X$ with $\text{dom } T \cap V \neq \emptyset$ the following holds: if $x \in V$ has the property that $\langle y^* - x^*, y - x \rangle \geq 0$ for all $y^* \in T (y)$, and all $y \in V$ then $x^* \in T (x)$.

**Corollary 34 (Convex Closure.)** Every maximal monotone mapping $T$ on a Banach space is maximal monotone locally. In particular, $\text{dom}(T)$ is convex.

**Proof.** We argue as follows. Fix $x, V$ and $x^*$ as in the definition of maximal monotonicity locally. We may select a closed convex set $C$ such that $x \in \text{int } C \subset V$ and $\text{dom } T \cap \text{int } C \neq \emptyset$. It follows from Theorem 31 that $T + N_C$ is maximal. Let $y^* \in T(y), n^* \in N_C(y), y \in V$ 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 final conclusion follows by results in [3] and earlier.

36


6 Appendix: Epi-limits

The give an outline of the bounded–weak* epi-limit–infimum and a proof of Theorem 40. This is a minor modification of the proof of Theorem 3.4 of [30]. Similar results may be found in [31] but are framed in a way that makes difficult the direct deduction of the result we require. The next result is a direct consequence of the Mrowka’s compactness theorem (see [1, Theorem 5.2.11]).

Proposition 35 1. Suppose $F : X \to Y$ is a multifunction between normed spaces. Then

$$\liminf_{v \to w} F(v) = \bigcap \{ C(w) \mid \exists \text{ a net } v_\beta \to w \text{ such that } \lim_{\beta} F(v_\beta) = C(w) \text{ exists} \}.$$  \hfill (72)

2. For a family of lower semicontinuous functions $f_v : X \to \mathbb{R}$ and a function $f : X \to \mathbb{R}$. Then we have that

$$e\text{-}\limsup_{v \to w} f_v \leq f$$ \hfill (73)

if and only if for all subnets $v_\beta \to w$ such that $e\text{-}\limsup_\beta f_{v_\beta} = e\text{-}\liminf_\beta f_{v_\beta}$ we have

$$e\text{-}\limsup_\beta f_{v_\beta} \leq f.$$ \hfill (74)

3. Consequently

$$e\text{-}\limsup_{v \to w} f_v = \sup \left\{ e\text{-}\limsup_\beta f_{v_\beta} \mid \exists \text{ sub-nets } v_\beta \to w \text{ s.t. } e\text{-}\lim_\beta f_{v_\beta} \text{ exists} \right\}.$$ \hfill (75)

Corollary 36 Let $f_v : X \to \mathbb{R}$ be a family of lower semicontinuous convex functions, $\{f_v\}_{v \in W}$ with $W$ a neighbourhood of $w$ and $f : X \to \mathbb{R}$ a lower semicontinuous, convex function. Then

$$\left( e\text{-}\limsup_{v \to w} f_v \right)^* \leq f$$

if and only if for all subnets $v_\beta \to w$ such that $e\text{-}\limsup_\beta f_{v_\beta} = e\text{-}\liminf_\beta f_{v_\beta}$ we have

$$\left( e\text{-}\limsup_\beta f_{v_\beta} \right)^* \leq f.$$
\begin{equation*}
eq -\infty \text{ implies } (e\limsup_{\beta} f_{v\beta})^* \equiv -\infty. \text{ On the other hand when } e\limsup_{v\rightarrow w} f_v \text{ is proper we have}
\end{equation*}
\begin{equation*}
(e\limsup_{v\rightarrow w} f_v)^* = e\limsup_{v\rightarrow w} f_v
\end{equation*}
and Proposition 35 implies \((e\limsup_{\beta} f_{v\beta})^* \leq f\) for all subnets \(v_\beta \rightarrow w\) such that \(e\limsup_{\beta} f_{v\beta} = e\liminf_{\beta} f_{v\beta}\).

Now suppose \((e\limsup_{v\rightarrow w} f_v)^* \beta (x) > f(x)\) for some \(x\) and so
\begin{equation*}
(e\limsup_{v\rightarrow w} f_v)^* \beta (x) > -\infty.
\end{equation*}
Then either \(e\limsup_{v\rightarrow w} f_v\) is proper or identically equal to \(+\infty\). In the former case properness implies \((e\limsup_{v\rightarrow w} f_v)^* = (e\limsup_{v\rightarrow w} f_v)\). Applying Proposition 35 part 3 we have the existence of subnets \(v_\beta \rightarrow w\) such that \(\{\text{epi } f_{v\beta}\}\) Kuratowski–Painlevé converges and for which \(e\limsup_{\beta} f_{v\beta} (x) > f(x)\). Because \(e\limsup_{\beta} f_{v\beta}\) is closed and convex we must have \(e\limsup_{\beta} f_{v\beta} \beta -\infty\) otherwise \(e\limsup_{\beta} f_{v\beta} \equiv -\infty\). Thus we conclude that
\begin{equation*}
(e\limsup_{v\rightarrow w} f_v)^* \beta (x) = e\limsup_{v\rightarrow w} f_{v\beta} (x) > f(x).
\end{equation*}
When \(e\limsup_{v\rightarrow w} f_v \equiv +\infty\) then invoking (75) also leads to the same conclusion.

We introduce the terminology of [30] for a family of lower semi–continuous proper function. Recall, that in order to show a convex set is bounded weak* closed it is sufficient to show it contains the limits of all bounded and weak* convergent nets taken from the set [14]. This observation motivates the following.

\textbf{Definition 37} Let \(\{f_v\}_{v \in W}\) be a family of functions on \(X\) and \(\{f^*_v\}_{v \in W}\) the family of conjugate functions on \(X^*\) (for a normed space \(X\)). We denote the bounded–weak* upper epi-limit (as \(v \rightarrow w\)) of \(\{f^*_v\}_{v \in W}\) by
\begin{equation*}
bw^*\limsup_{v \rightarrow w} \text{epi } f^*_v := \{(x^*, \alpha) \in X^* \times \mathbb{R} \mid \exists \text{ nets } v_\beta \rightarrow w; (y^*_\beta, \alpha_\beta) \in \text{epi } f^*_v \beta \text{ such that } \alpha_\beta \rightarrow \alpha; y^*_\beta \text{ norm bounded; } y^*_\beta \overset{w}{\rightarrow} x^*\}.
\end{equation*}

The above closely resembles the limit–superior of epigraphs, relative to the bounded–weak* topology on \(X^*\) (hence the terminology). Clearly this set recedes to \(+\infty\) in the vertical direction and so resembles the epigraph of some function. This prompts us to define

\textbf{Definition 38} For \(x^* \in X^*\),
\begin{equation*}
(bw^*e\liminf_{v \rightarrow w} f^*_v)(x^*) := \inf \{\alpha \in \mathbb{R} \mid (x^*, \alpha) \in bw^*\limsup_{v \rightarrow w} \text{epi } f^*_v\}. \quad (76)
\end{equation*}
It then follows that
\[
\text{epi}_\mathcal{A} (bw^*\text{-}e\text{-}\lim\inf_{v\to w} f^*_v) \subseteq bw^*\text{-}\lim\sup_{v\to w} f^*_v \subseteq \text{epi} (bw^*\text{-}e\text{-}\lim\inf_{v\to w} f^*_v).
\]
(77)

Thus \(bw^*\text{-}e\text{-}\lim\inf_{v\to w} f^*_v\) is essentially a variational limit in the sense of use by Aubin, Rockafellar and Wets. Analogous definitions can be made for nets \(\{f_\gamma\}_{\gamma \in I}\) of functions i.e.
\[
bw^*\text{-}\lim\sup_{\gamma} \text{epi} f^*_\gamma := \{(x^*, \alpha) \in X^* \times \mathbb{R} \mid \exists \text{ subnet } \gamma_\beta : (y^*_\beta, \alpha_\beta) \in \text{epi } f^*_\gamma \text{ such that } \alpha_\beta \to \alpha; y^*_\beta \text{ norm bounded; } y^*_\beta \overset{w^*}{\to} x^*\}.
\]
with \((bw^*\text{-}e\text{-}\lim\inf_{\gamma} f^*_\gamma)(x^*)\) defined as in (76).

We now state Lemma 3.3 of [30].

**Proposition 39** Let \(X\) be a normed space, and \(\{f_\beta\}_{\beta \in I}\) a net of proper closed convex extended–real–valued functions on \(X\). Suppose also that the strong epi–limit of \(\{f_\beta\}_{\beta \in I}\) exists. Also, assume that either:

1. this epi–limit takes a finite value somewhere,
2. or \(bw^*\text{-}e\text{-}\lim\inf_{\beta} f^*_\beta\) is not identically \(+\infty\).

Then
\[
e\text{-}\lim\sup_{\beta} f \leq (bw^*\text{-}e\text{-}\lim\inf_{\beta} f^*_\beta)^*.
\]
(78)

The proof in [30] shows that the inequality (78) holds at each \(x \in \text{dom } [e\text{-}\lim\sup_{\beta} f(\cdot)]\) without any of the additional assumptions stated in Theorem 39. We now weaken the assumption for which (78) holds outside this domain. This next result is proved using an adoption of the proof of Lemma 3.4 of [30].

**Theorem 40** Let \(X\) be a normed space, \(W\) a topological space; let \(\{f_v\}_{v \in W}\) be a family of proper closed convex extended–real–valued functions on \(X\). Suppose in addition that \(e\text{-}\lim\sup_{v \to w} f_v > -\infty\). Then
\[
(e\text{-}\lim\sup_{v \to w} f_v)^{**} \leq (bw^*\text{-}e\text{-}\lim\inf_{v \to w} f^*_v)^*.
\]
(79)

**Proof.** Recall that the epi–limit–supremum satisfies
\[
\text{epi} \left( e\text{-}\lim\sup_{\gamma} f_{v_\beta, \gamma} \right) = \lim\inf_{\gamma} \text{epi } f_{v_\beta, \gamma}
\]
and hence is a closed convex function. We only have to consider nets \(v_\beta \to w\) so that there is a strongly epi–convergent subnet \(f_{v_\beta, \gamma}\) with \(e\text{-}\lim\sup_{\gamma} f_{v_\beta, \gamma}\) finite at some point. Indeed take any \(v_\beta \to w\) so that there is a strongly epi–convergent subnet \(\{f_{v_\beta, \gamma}\}\) then if we suppose that for all \(x \in \text{dom } (e\text{-}\lim\sup_{\gamma} f_{v_\gamma})\)
we have $e\text{-lim sup} \gamma f_{v,\beta,\gamma} (x) = -\infty$ (otherwise a finite value is attained) then we find that for all $x \in \text{dom} \left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)$

$$e\text{-lim sup} \gamma f_{v,\beta,\gamma} (x) = \left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)^{**} (x) = -\infty.$$  

For $x \notin \text{dom} \left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)$ we have $\left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)^{**} (x) = -\infty$ while $e\text{-lim sup} \gamma f_{v,\beta,\gamma} (x) = +\infty$. Consequently we have

$$\left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)^{**} \leq (bw^*-e\text{-lim inf}_{v \to w} f_v^*)^*$$

holding on all of $X$. Then we need only to appeal to Proposition 35 and Lemma 39 which considering only subnets $v_\beta \to w$ for which the there is a strongly epi–convergent subnet $f_{v,\beta,\gamma}$ with $e\text{-lim sup} \gamma f_{v,\beta,\gamma}$ finite at some point. Then by Proposition 39 (applied to $f_{v,\beta,\gamma}$),

$$\left( e\text{-lim sup} \gamma f_{v,\beta,\gamma} \right)^{**} \leq e\text{-lim sup} \gamma f_{v,\beta,\gamma} \leq (bw^*-e\text{-lim inf}_{v \to w} f_v^*)^*$$

$$\leq (bw^*-e\text{-lim inf}_{v \to w} f_v^*)^*.$$  

Since this inequality holds for all convergent subnets $v_\beta \to w$ with $f_{v,\beta,\gamma}$ strongly epi–convergent, by invoking Corollary 36 we conclude that (79) holds.

Acknowledgment: The authors would like to thank C. Zălinescu for valuable comments and criticism.

References


