MAXIMALITY OF SUMS OF TWO MAXIMAL MONOTONE OPERATORS

JONATHAN M. BORWEIN, FRSC

Abstract. We use methods from convex analysis convex, relying on an ingenious function of Simon Fitzpatrick, to prove maximality of the sum of two maximal monotone operators on reflexive Banach space under weak transversality conditions.

1. Introduction and Preliminaries

The central result of this paper, Theorem 5, marries recent work by Simons and Zalinescu [16] with additional convex analysis to provide an accessible short proof of the maximality of the sum of two maximal monotone operators.

Recall that the *domain* of an extended valued convex function, denoted $\mathrm{dom}\,(f)$, is the set of points with value less than $+\infty$, and that a point s is in the *core* of a set S (denoted by $s \in \mathrm{core}\, S$) provided that s lies in S and $X = \bigcup_{\lambda>0} \lambda(S-s)$. For a concave function g, we use $\mathrm{dom}\, g = \mathrm{dom}\,(-g)$. Recall that $x^* \in X^*$ is a *subgradient* of $f : X \to (-\infty, +\infty]$ at $x \in \mathrm{dom}\, f$ provided that $f(y) - f(x) \geq \langle x^*, y - x\rangle$. The set of all subgradients of f at x is called the *subderivative* or *subdifferential* of f at x and is denoted $\partial f(x)$. We use the convention that $\partial f(x) = \emptyset$ for $x \notin \mathrm{dom}\, f$. We shall need the *indicator* function $\iota_C(x)$ which is zero for x in C and $+\infty$ otherwise, the *Fenchel conjugate* $f^*(x^*) := \sup_x\{\langle x, x^*\rangle - f(x)\}$ and the *infimal convolution* $f^*\Box\frac{1}{2}\|\cdot\|_*^2(x^*) := \inf\{f^*(y^*) + \frac{1}{2}\|z^*\|_*^2 \colon x^* = y^* + z^*\}$. When f is convex and closed and x is in the domain of f, $x^* \in \partial f(x)$ exactly when $f(x) + f^*(x^*) = \langle x, x^*\rangle$.

We say a multifunction $F : X \mapsto 2^{X^*}$ is *monotone* provided that for any $x, y \in X$, $x^* \in F(x)$ and $y^* \in F(y)$,

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

and we say that T is *maximal monotone* if its graph is not properly included in any other monotone graph. The subdifferential of a convex lower semicontinuous (lsc) function on a Banach space is a typical example of a maximal monotone multifunction (see [4, 6, 13] wherein other notation and usage may be also followed up). Indeed we reserve the notation J for the duality map

$$J(x) := \frac{1}{2}\,\partial\|x\|^2 = \{x^* \in X^* \colon \|x\|^2 = \|x^*\|^2 = \langle x, x^*\rangle\}.$$

Further applications and a significantly more extended discussion of the techniques in this note can be found in [1].

Date: August 31, 2005.

1991 *Mathematics Subject Classification.* 47H05, 46N10, 46A22.

Key words and phrases. monotone operators, convex analysis, sandwich theorem, Fenchel duality, sum theorem.

Research was supported by NSERC and by the Canada Research Chair Program.

Proposition 1. [4, 6, 13] *For a closed convex function f, let $f_J := f + \frac{1}{2}\|\cdot\|^2$. Then $f_J^* = \left(f + \frac{1}{2}\|\cdot\|^2\right)^* = f^* \Box \frac{1}{2}\|\cdot\|_*^2$ is everywhere continuous. Also*

$$v^* \in \partial f(v) + J(v) \Leftrightarrow f_J^*(v^*) + f_J(v) - \langle v, v^* \rangle \leq 0.$$

For any monotone mapping T, we associate the *Fitzpatrick function* introduced by Simon Fitzpatrick in [8] but then neglected for many years until re-popularized in papers by Penot [10], Buracik-Svaiter [7], and others. Some more of the related history may be found in [1]. Fitzpatrick's function is

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

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

Proposition 2. [8, 6] *For a maximal monotone operator T*

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

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

We recall the version of the Hahn-Banach theorem we need:

Theorem 3. (Hahn-Banach Sandwich, [4, 6, 13]) *Suppose f and $-g$ are proper extended real-valued lsc convex on a Banach space X and that $f(x) \geq g(x)$, for all x in X. Assume that*

$$0 \in \operatorname{core}\ (\operatorname{dom}(f) - \operatorname{dom}(g)).$$

Then there is a continuous linear function λ such that

$$f(x) - g(y) \geq \langle \lambda, x - y \rangle,$$

for all $x \in \operatorname{dom} f, y \in \operatorname{dom} -g$ in X.

Proof. The value function $h(u) := \inf_X f(x) - g(x-u)$ is convex. It is continuous at 0—indeed the constraint qualification and semi-continuity of the data force h to be bounded above around zero—by a Baire category type argument [17, 14, 6]. Hence there is some $-\lambda \in \partial h(0)$. This provides the linear part of the asserted affine separator. Indeed, we have

$$f(x) - g(u - x) \geq h(u) - h(0) \geq \lambda(u - 0),$$

as required. □

The next result, implicit in the literature [14], avoids needing to renorm a reflexive space to have a single-valued duality map with a single-valued inverse.

Proposition 4. [14, 11, 12, 6] *A monotone multifunction T is maximal if and only if the mapping $T(\cdot + w) + J$ is surjective for all w in X.* [*When J and J^{-1} are both single valued, a monotone mapping T is maximal if and only if $T + J$ is surjective.*]

Proof. We prove the 'if'. The 'only if' is completed in Corollary 7. Assume (w, w^*) is monotonically related to the graph of T. By hypothesis, we may solve $w^* \in T(x + w) + J(x)$. Thus $w^* = t^* + j^*$ where $t^* \in T(x + w), j^* \in J(x)$.

$$0 \leq \langle w^* - t^*, w - (w + x) \rangle = -\langle w^* - t^*, x \rangle = -\langle j^*, x \rangle = -\|x\|^2 \leq 0.$$

Hence $x = 0, j^* = 0$ and we are done.

The refined equivalence when J, J^{-1} are single-valued may be found in [14, Thm 10.6]. □

2. The Main Result

We now prove our asserted result—whose proof originally very hard and due to Rockafellar [12]—has been revisited over many years culminating in [14, 7, 10, 16, 6] among others. The proof we give is perhaps the first to avoid using either renorming [12] or some preliminary minimax arguments [14]:

Theorem 5. *Let X be any reflexive space with given norm. Let T be maximal monotone and f closed and convex. Suppose that*

$$0 \in \operatorname{core}\{\operatorname{conv}\operatorname{dom}(T) - \operatorname{conv}\operatorname{dom}(\partial f)\}.$$

Then

(a) $\partial f + T + J$ *is surjective;*
(b) $\partial f + T$ *is maximal monotone;*
(c) ∂f *is maximal monotone.*

Proof. (a) As in [16], we consider the Fitzpatrick function $F_T(x, x^*)$ and further introduce $f_J(x) := f(x) + 1/2\|x\|^2$. Let $G(x, x^*) := -f_J(x) - f_J^*(-x^*)$. Observe that

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

pointwise thanks to Proposition 2, and the *Fenchel-Young inequality*: for any function $f(x) + f^*(x^*) \geq \langle x, x^*\rangle, \forall x, x^*$. Now, the (CQ)

$$0 \in \operatorname{core}\{\operatorname{conv}\operatorname{dom}(T) - \operatorname{conv}\operatorname{dom}(\partial f)\}$$

assures the *Sandwich theorem* applies. Indeed, by Proposition 1, f_J^* is everywhere finite and in consequence zero is in the core of $\operatorname{dom} F_T - \operatorname{dom} G$.

Then there are $w \in X$ and $w^* \in X^*$ such that

$$F_T(x, x^*) - G(z, z^*) \geq w(x^* - z^*) + w^*(x - z) \tag{1}$$

so that for all $x^* \in T(x), x \in \operatorname{dom}(T)$ and for all z^*, z we have

$$\langle x^* - w^*, x - w\rangle + [f_J(z) + f_J^*(-z^*) + \langle z, z^*\rangle] \geq \langle w^* - z^*, w - z\rangle.$$

Now use the fact that $-w^* \in \operatorname{dom}(\partial f_J^*)$, by Proposition 1, to deduce that for some v, $-w^* \in \partial f_J(v)$ and so

$$\langle x^* - w^*, x - w\rangle + [f_J(v) + f_J^*(-w^*) + \langle v, w^*\rangle] \geq \langle w^* - w^*, w - v\rangle = 0.$$

The second term on the left is zero and so $w^* \in T(w)$ by maximality. Substitution of $x = w$ and $x^* = w^*$ in (1), and rearranging yields

$$\langle w^*, w\rangle + \{\langle -z^*, w\rangle - f_J^*(-z^*)\} + \{\langle z, -w^*\rangle - f_J(z)\} \leq 0,$$

for all z, z^*. Taking the supremum over z and z^* produces $\langle w^*, w\rangle + f_J(w) + f_J^*(-w^*) \leq 0$. This shows $-w^* \in \partial f_J(w) = \partial f(w) + J(w)$ on using the sum formula for subgradients, implicit in Proposition 1.

Thus, $0 \in (T + \partial f_J)(w)$ and, since all range translations of $T + \partial f$ may be used, $\partial f + T + J$ is surjective which completes (a). Additionally, since all domain translations may be used, $\partial f + T$ is maximal by the easy part of Proposition 4, which yields (b).

Finally, setting $T \equiv 0$ we recover the reflexive case of the maximality for a lsc convex function ∂f which is (c). □

Note that we have exploited the beautiful inequality

$$F_T(x, x^*) + f(x) + f^*(-x^*) \geq 0, \qquad \forall x \in X, x^* \in X^*, \tag{2}$$

valid for *any* maximal monotone T and *any* convex function f.

3. Some Corollaries

We first recover the so called *Brezis-Attouche theorem*:

Corollary 6. [14] *The sum of two maximal monotone operators* T_1 *and* T_2 *is maximal monotone if* $0 \in \operatorname{core}[\operatorname{conv}\operatorname{dom}(T_1) - \operatorname{conv}\operatorname{dom}(T_2)]$.

Proof. Theorem 5 applies to the maximal monotone mapping $T(z) := (T_1(x), T_2(y))$ and the indicator function $f(x, y) = \iota_{\{x=y\}}$. Finally, check that the given transversality condition implies the needed (CQ). We obtain that $T + J_{X \otimes X} + \partial\iota_{\{x=y\}}$ is surjective. Thus, so is $T_1 + T_2 + 2J$ and we are done. □

We next recover the *Rockafellar-Minty surjectivity theorem*:

Corollary 7. *A maximal monotone on a reflexive space has* $\operatorname{range}(T + J) = X^*$.

Proof. Let $f \equiv 0$ in Theorem 5. □

Recall that T is *coercive on* C if $\inf_{y^* \in (T+\partial\iota_C)(y)} \langle y, y^*\rangle / \|y\| \to \infty$ as $y \in C$ goes to infinity in norm, with the convention that $\inf \emptyset = +\infty$. A *variational inequality* requests a solution $y \in C$ and $y^* \in T(y)$ to

$$\langle y^*, x - y\rangle \geq 0 \qquad \forall x \in C.$$

We denote the variational inequality by $V(T; C)$.

Corollary 8. *Suppose* T *is maximal monotone on a reflexive Banach space and is coercive on the closed convex set* C. *Suppose also that* $0 \in \operatorname{core}(C - \operatorname{conv}\operatorname{dom}(T))$. *Then* $V(T, C)$ *has a solution.*

Proof. Let $f := \iota_C$, the indicator function. For $n = 1, 2, 3 \cdots$, let $T_n := T + J/n$. We solve

$$0 \in (T_n + \partial\iota_C)(y_n) = (T + \partial\iota_C) + \frac{1}{n} J(y_n) \tag{3}$$

and take limits as n goes to infinity. More precisely, we observe that using our key Theorem 5, we find y_n in C, and $y_n^* \in (T + \partial\iota_C)(y_n), j_n^* \in J(y_n)/n$ with $y_n^* = -j_n^*$. Then

$$\langle y_n^*, y_n\rangle = -\frac{1}{n}\langle j_n^*, y_n\rangle = -\frac{1}{n}\|y_n\|^2 \leq 0$$

so coercivity of $T + \partial\iota_C$ implies that $\|y_n\|$ remains bounded and so $j_n^* \to 0$. On taking a subsequence we may assume $y_n \rightharpoonup y$. Since $T + \partial\iota_C$ is maximal monotone (again by Theorem 5), it is demi-closed [6]. It follows that $0 \in (T + \partial\iota_C)(y)$ as required. □

Letting $C = X$ we deduce:

Corollary 9. *Every coercive maximal monotone multifunction on a Banach space is surjective if (and only if) the space is reflexive.*

Proof. To complete the proof we recall that, by James' theorem, surjectivity of J is equivalent to reflexivity of the corresponding space. □

Details of these and other Corollaries are to be found in [1].

Finally, warm thanks are due to Jean-Paul Penot and to Jim Zhu for their very helpful and detailed comments about this note.

References

[1] Borwein J. M., "Maximal montonicity via convex analysis," in preparation for Fitzpatrick memorial issue of *J. Convex Analysis*, 2005.

[2] Borwein J. M. and Fitzpatrick, S. , "Local boundedness of monotone operators under minimal hypotheses," *Bull. Australian Math. Soc.*, **39** (1988), 339–441.

[3] J.M. Borwein, S. Fitzpatrick and J. Vanderwerff, "Examples of convex functions and classifications of normed spaces," *Journal of Convex Analysis*, **1** (1994), 61–73.

[4] Borwein J. M. & Lewis A. S., *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer-Verlage, New York, 2000.

[5] Borwein J. M. & Zhu Q. J., "A survey of subdifferential calculus with applications," Nonlinear Analysis, TMA, **38** (1999), 687-773.

[6] Borwein J. M. & Zhu Q. J., *Techniques of Variational Analysis: an Introduction*, CMS Books, Springer-Verlag, 2005.

[7] Burachik, R.S., & Svaiter, B.F., Maximal monotonicity, conjugation and the duality product," *Proc AMS,* **131** (2003), 2379–2383.

[8] Fitzpatrick, S. "Representing monotone operators by convex functions," Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), 59–65, Proc. Centre Math. Anal. Austral. Nat. Univ., 20, Austral. Nat. Univ., Canberra, 1988.

[9] Borwein J. M. & Zhu Q. J., *A survey of subdifferential calculus with applications*, Nonlinear Analysis, TMA, **38** (1999), 687-773.

[10] Penot J.P., "The relevance of convex analysisi for the astudy of montonicity," *Nonlinear Anal, Theory, Methods, Appl.*, **58** (2004),855–871.

[11] Rockafellar R. T., "Local boundedness of nonlinear monotone operators," *Mich. Math. J.*, **16** (1969), 397–407.

[12] Rockafellar R. T., "On the maximality of sums of nonlinear monotone operators," *Trans. Amer. Math. Soc.*, **149** (1970), 75–88.

[13] Rockafellar R. T., *Convex Analysis*, Princeton University Press, Princeton, N.J., 1970.

[14] Simons, S., *Minimax and Monotonicity*, Lecture Notes in Mathematics, **1693**, Springer-Verlag,Berlin, 1998.

[15] Simons, S., "A new version of the Hahn-Banach theorem," preprint available at: `www.math.ucsb.edu/~simons/preprints`, (2003).

[16] Simons, S. and C. Zalinescu, "A new proof for Rockafellar's characterization of maximal monotone operators," *Proc. Amer. Math. Soc.* **132** (2004), 2969–2972.

[17] Verona, A. and Verona, M. E., "Regular maximal monotone operators and the sum rule," *J. Convex Analysis*, **7** (2000), 115–128.

Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada, e-mail: jborwein@cs.dal.ca