Brøndsted-Rockafellar property of subdifferentials of prox-bounded functions

Marc Lassonde
Université des Antilles et de la Guyane

Playa Blanca, Tongoy, Chile
Everywhere $X$ is a Banach space. A set-valued operator $T : X \rightrightarrows X^*$, or graph $T \subset X \times X^*$, is **monotone** provided
\[
\langle y^* - x^*, y - x \rangle \geq 0, \quad \forall (x, x^*), (y, y^*) \in T,
\]
and **maximal monotone** provided it is monotone and not properly contained in another monotone operator.

The **subdifferential** $\partial f : X \rightrightarrows X^*$ of a convex $f : X \to ]-\infty, +\infty]$ is
\[
\partial f(x) := \left\{ x^* \in X^* : \langle x^*, y - x \rangle + f(x) \leq f(y), \forall y \in X \right\},
\]
and the **duality operator** $J : X \rightrightarrows X^*$ is
\[
J(x) := \left\{ x^* \in X^* : \langle x^*, x \rangle = \|x\|^2 = \|x^*\|^2 \right\}.
\]
It is easily verified that $J(x) = \partial j(x)$ where $j(x) = (1/2)\|x\|^2$.

**Theorem (Rockafellar, 1970)** Let $X$ be a Banach space. The subdifferential $\partial f$ of a proper convex lower semicontinuous function $f : X \to ]-\infty, +\infty]$ is maximal monotone.
PROOF WHEN $X = H$ IS HILBERT (taken from Brezis, 1973)

By Hahn-Banach, $f \geq \ell + \alpha$ for some $\ell \in X^*$ and $\alpha \in \mathbb{R}$, and $j + \ell$ is coercive $(j(x) + \ell(x) = (1/2)\|x\|^2 + \ell(x) \to +\infty$ as $\|x\| \to +\infty)$, so

\[ f + j \text{ is coercive.} \]

Hence $f + j$ attains its minimum at some $\bar{x} \in H$, so $0 \in \partial(f + j)(\bar{x})$.

Since $\partial j = \nabla j = I$ (identity on $H$), we readily get $0 \in (\partial f + I)(\bar{x})$, so $0 \in R(\partial f + I)$. We conclude that

\[ X^* = R(\partial f + I). \]

This is easily seen to imply that $\partial f$ is maximal monotone.

(This is the elementary part in Minty’s characterization of maximal monotonicity (1962).)
PROOF IN THE GENERAL BANACH CASE: STEP 1

Claim:

\[ 0 \in \overline{R(\partial f + J)}. \]  \hspace{1cm} (1)

First, \( f \geq \ell + \alpha \) for some \( \ell \in X^* \) and \( \alpha \in \mathbb{R} \), and \( j + \ell \) bounded below, so \( f + j \) is bounded below.

Next, let \( \varepsilon > 0 \) arbitrary and let \( y_\varepsilon \in \text{dom } f \) such that

\[ (f + j)(y_\varepsilon) \leq (f + j)(y) + \varepsilon^2, \forall y \in X. \]

By Brøndsted-Rockafellar approximation theorem (1965),

\[ \exists x_\varepsilon^* \in X^* \text{ with } \|x_\varepsilon^*\| \leq \varepsilon \text{ and } z_\varepsilon \in X \text{ such that } x_\varepsilon^* \in \partial(f + j)(z_\varepsilon). \]

By Rockafellar’s sum rule (1966), \( x_\varepsilon^* \in \partial f(z_\varepsilon) + J(z_\varepsilon). \)

Conclusion: \( \exists x_\varepsilon^* \in R(\partial f + J) \text{ with } \|x_\varepsilon^*\| \leq \varepsilon, \) proving the claim.
**PROOF: STEP 2**

Let \((x, x^*) \in X \times X^*\) such that
\[
\langle y^* - x^*, y - x \rangle \geq 0, \quad \forall (y, y^*) \in \partial f. \tag{2}
\]

Applying [1] to \(f(x + .) - x^*\), we get
\[
x^* \in R(\partial f(x + .) + J).
\]

Thus, there are \((x_n^*) \subset X^*\) with \(x_n^* \to x^*\) and \((h_n) \subset X\) such that
\[
x_n^* \in \partial f(x + h_n) + J(h_n).
\]

Thus, there are \((x_n^*) \subset X^*\) with \(x_n^* \to x^*\) and \((h_n) \subset X\) such that
\[
y_n^* \in \partial f(x + h_n) \quad \text{and} \quad x_n^* - y_n^* \in J(h_n).
\]

By definition of \(J\), we have
\[
\langle x_n^* - y_n^*, h_n \rangle = \|x_n^* - y_n^*\|^2 = \|h_n\|^2. \tag{3}
\]

From [2] and \(y_n^* \in \partial f(x + h_n)\), we get \(\langle x^* - y_n^*, x + h_n - x \rangle \leq 0\), so
\[
\|h_n\|^2 = \langle x_n^* - x^*, h_n \rangle + \langle x^* - y_n^*, x + h_n - x \rangle \leq \langle x_n^* - x^*, h_n \rangle \leq \|x_n^* - x^*\|\|h_n\|.
\]

Hence, \(h_n \to 0\), so, by [3], \(\|x_n^* - y_n^*\| \to 0\), therefore \(y_n^* \to x^*\). Since \(\partial f\) has closed graph and \(y_n^* \in \partial f(x + h_n)\), we conclude that \(x^* \in \partial f(x)\).
OTHER PROOFS OF MAXIMALITY OF $\partial f$ FOR CONVEX $f$

1/ $f$ everywhere finite and continuous:
- Minty (1964), Phelps (1989), using mean value theorem and link between subderivative and subdifferential

2/ $f$ lsc, $X$ Hilbert:
- Moreau (1965), via prox functions and duality theory,
- Brezis (1973), showing directly that $\partial f + I$ is onto

3/ $f$ lsc, $X$ Banach: all proofs use a variational principle and another tool
- Rockafellar (1970): continuity of $(f + j)^*$ in $X^*$ and link between $(\partial f)^{-1}$ and $\partial f^*$ in $X^{**} \times X^*$,
- Taylor (1973) and Borwein (1982): subderivative mean value inequality and link between subderivative and subdifferential,
- Zagrodny (1988?), Simons (1991), Luc (1993), etc: subdifferential mean value inequality,
- Thibault (1999): limiting convex subdifferential calculus,
BEYOND THE CONVEX CASE: MAIN TOOLS

Let be given a 'subdifferential' \( \partial \) that associates a subset \( \partial f(x) \subset X^* \) to each \( x \in X \) and each function \( f \) on \( X \) so that \( \partial f(x) \) coincides with the convex subdifferential when \( f \) is convex.

The two main tools in the convex situation were:

- \textit{Brøndsted-Rockafellar's approximation theorem} (1965)
- \textit{Rockafellar's subdifferential sum rule} (1966).

They will be respectively replaced by:

\begin{center}
\textbf{Ekeland Variational Principle} (1974). For any lsc function \( f \) on \( X \), \( \bar{x} \in \text{dom } f \) and \( \varepsilon > 0 \) such that \( f(\bar{x}) \leq \inf f(X) + \varepsilon \), and for any \( \lambda > 0 \), there is \( x_\lambda \in X \) s.t. \( \|x_\lambda - \bar{x}\| \leq \lambda \), \( f(x_\lambda) \leq f(\bar{x}) \), and
  \[ x \mapsto f(x) + (\varepsilon/\lambda)\|x - x_\lambda\| \] has a minimum at \( x_\lambda \).
\end{center}

\begin{center}
\textbf{Subdifferential Separation Principle}. For any lsc functions \( f, \varphi \) on \( X \) with \( \varphi \) convex Lipschitz near \( \bar{x} \in \text{dom } f \cap \text{dom } \varphi \),
  \[ f + \varphi \] has a local minimum at \( \bar{x} \implies 0 \in \partial f(\bar{x}) + \partial \varphi(\bar{x}) \).
\end{center}
SUBDIFFERENTIALS SATISFYING THE SEPARATION PRINCIPLE

The main examples of pairs \((X, \partial)\) for which the Subdifferential Separation Principle holds are:

- the Clarke subdifferential \(\partial_C\) in arbitrary Banach spaces,
- the limiting Fréchet subdifferential \(\hat{\partial}_F\) in Asplund spaces,
- the limiting Hadamard subdifferential \(\hat{\partial}_H\) in separable spaces,
- the limiting proximal subdifferential \(\hat{\partial}_P\) in Hilbert spaces.

For more details, see, e.g., Jules-Lassonde (2013, 2013b).
COMBINING THE TOOLS

Set \( \text{dom } f^* = \{ x^* \in X^* : \inf(f - x^*)(X) > -\infty \} \).

**Proposition** Let \( X \) Banach, \( f : X \to ]-\infty, +\infty[ \) proper lsc, \( \varphi : X \to \mathbb{R} \) convex loc. Lispchitz. Then, \( \text{dom } (f + \varphi)^* \subset \text{cl } (R(\partial f + \partial \varphi)) \).

**Proof.** Let \( x^* \in \text{dom } (f + \varphi)^* \) and let \( \varepsilon > 0 \). There is \( \bar{x} \in X \) s.t.

\[
(f + \varphi - x^*)(\bar{x}) \leq \inf(f + \varphi - x^*)(X) + \varepsilon^2,
\]

so, by Ekeland’s variational principle, there is \( x_{\varepsilon} \in X \) such that

\[
x \mapsto f(x) + \varphi(x) + \langle -x^*, x \rangle + \varepsilon \| x - x_{\varepsilon} \| \text{ attains its minimum at } x_{\varepsilon}.
\]

Now, applying the Separation Principle with the convex locally Lipschitz \( \psi : x \mapsto \varphi(x) + \langle -x^*, x \rangle + \varepsilon \| x - x_{\varepsilon} \| \) we obtain \( x_{\varepsilon}^* \in \partial f(x_{\varepsilon}) \) such that

\[-x_{\varepsilon}^* \in \partial \psi(x_{\varepsilon}) = \partial \varphi(x_{\varepsilon}) - x^* + \varepsilon B_{X^*}.
\]

So, there is \( y_{\varepsilon}^* \in \partial \varphi(x_{\varepsilon}) \) such that

\[\| x^* - y_{\varepsilon}^* - x_{\varepsilon}^* \| \leq \varepsilon.\]

Thus, for every \( \varepsilon > 0 \) the ball \( B(x^*, \varepsilon) \) contains \( x_{\varepsilon}^* + y_{\varepsilon}^* \in \partial f(x_{\varepsilon}) + \partial \varphi(x_{\varepsilon}) \subset R(\partial f + \partial \varphi) \). This means that

\( x^* \in \text{cl } (R(\partial f + \partial \varphi)) \).

The case \( \varphi = 0 \) and \( f = \delta_C \) with \( C \) nonempty closed convex set says that the set \( R(\partial \delta_C) \) of functionals in \( X^* \) that attain their supremum on \( C \) is dense in the set \( \text{dom } \delta_{\varepsilon}^* \) of all those functionals which are bounded above on \( C \) (Bishop-Phelps).
**PROX-BOUNDED FUNCTIONS**

A function $f$ is called *prox-bounded* if there exists $\lambda > 0$ such that the function $f + \lambda j$ is bounded below; the infimum $\lambda_f$ of the set of all such $\lambda$ is called the *threshold* of prox-boundedness for $f$:

$$\lambda_f := \inf\{\lambda > 0 : \inf(f + \lambda j) > -\infty\}.$$ 

Any convex lsc function $g$ is prox-bounded with threshold $\lambda_g = 0$, the sum $f + g$ of a prox-bounded $f$ and of a convex lsc $g$ is prox-bounded with $\lambda_{f+g} \leq \lambda_f$, for every $x^* \in X^*$, $\lambda_{f+x^*} = \lambda_f$, and for every $x \in X$, $f(x + .) + \lambda j$ is bounded below for any $\lambda > \lambda_f$ (see Rockafellar-Wets book (1998)).

Consequence: if $f$ is prox-bounded, then for every $\lambda > \lambda_f$,

$$\forall x \in X, \ \text{dom}(f(x + .) + \lambda j)^* = X^*.$$ 

From this and the previous result we get:

**Proposition** Let $X$ Banach and let $f : X \to ]-\infty, +\infty]$ be lsc and prox-bounded with threshold $\lambda_f$. Then, for every $\lambda > \lambda_f$,

$$\forall x \in X, \ \text{cl}(R(\partial f(x + .) + \lambda J)) = X^*.$$
GOING FURTHER: MONOTONE ABSORPTION

Given $T : X \ni X^*$, or $T \subset X \times X^*$, and $\varepsilon \geq 0$, we let

$$T^\varepsilon := \{(x, x^*) \in X \times X^* : \langle y^* - x^*, y - x \rangle \geq -\varepsilon, \ \forall (y, y^*) \in T\}$$

be the set of pairs $(x, x^*)$ $\varepsilon$-monotonically related to $T$.

An operator $T$ is monotone provided $T \subset T^0$ and monotone maximal provided $T = T^0$.

A non necessarily monotone operator $T$ is declared to be monotone absorbing provided $T^0 \subset \overline{T}$ (norm-closure).

A non necessarily monotone operator $T$ is declared to be widely monotone absorbing with threshold $\lambda_T \geq 0$ provided for every $\lambda > \lambda_T$ one has

$$\forall \varepsilon \geq 0, \ T^\varepsilon \subset \left( T + \sqrt{\lambda^{-1}\varepsilon} B_X \times \sqrt{\lambda \varepsilon} B_{X^*} \right).$$

Equivalently: $\forall \varepsilon \geq 0$, $(x, x^*) \in T^\varepsilon \Rightarrow$

$$\exists (x_n, x_n^*) \subset T : \lim_n \|x - x_n\| \leq \sqrt{\lambda^{-1}\varepsilon} \text{ and } \lim_n \|x^* - x_n^*\| \leq \sqrt{\lambda \varepsilon}.$$
**SUFFICIENT CONDITION FOR WIDE MONOTONE ABSORPTION**

**Proposition** Let \( T : X \to X^* \) and \( \lambda > 0 \). Assume that

\[
\forall x \in X, \ \text{cl} (R(T(x + .) + \lambda J)) = X^*. \tag{4}
\]

Then:

\[
\forall \varepsilon \geq 0, \ T^\varepsilon \subset \text{cl} \left( T + \sqrt{\lambda^{-1}} \varepsilon B_X \times \sqrt{\lambda} \varepsilon B_{X^*} \right). \tag{5}
\]

**Proof.** Let \( \varepsilon \geq 0 \) and let \((x, x^*) \in T^\varepsilon \). Since \( T(x + .) + \lambda J \) has a dense range, we can find \((x_n^*) \subset X^* \) with \( x_n^* \to x^* \) and \((y_n) \subset X \) such that \( x_n^* \in T(x + y_n) + \lambda J y_n \). Let \((y_n^*) \subset X^* \) such that

\[
y_n^* \in T(x + y_n) \quad \text{and} \quad x_n^* - y_n^* \in \lambda J y_n.
\]

By definition of \( J \), we have

\[
\lambda^{-1} \langle x_n^* - y_n^*, y_n \rangle = \| \lambda^{-1} (x_n^* - y_n^*) \|^2 = \| y_n \|^2. \tag{6}
\]

But \((x_n^*) \subset T^\varepsilon x \) and \((y_n^*) \subset T(x + y_n) \), so \( \langle x^* - y_n^*, y_n \rangle \leq \varepsilon \), hence

\[
\lambda \| y_n \|^2 = \langle x_n^* - x^*, y_n \rangle + \langle x^* - y_n^*, y_n \rangle \leq \langle x_n^* - x^*, y_n \rangle + \varepsilon \leq \| x_n^* - x^* \| \| y_n \| \varepsilon.
\]
Therefore, $\lambda \| y_n \|^2 - \| x_n^* - x^* \| \| y_n \| - \varepsilon \leq 0$, so we must have

$$\| y_n \| \leq \left( \| x_n^* - x^* \| + \sqrt{\| x_n^* - x^* \|^2 + 4\varepsilon \lambda} \right) / 2\lambda. \quad (7)$$

From (7) we derive that $\limsup_n \| y_n \| \leq \sqrt{\lambda^{-1} \varepsilon}$, so, by (6),

$$\limsup_n \| x_n^* - y_n^* \| = \limsup_n \lambda \| y_n \| \leq \sqrt{\lambda \varepsilon}.$$ 

In conclusion we have $(x + y_n, y_n^*) \in T$ with

$$\limsup_n \| x - (x + y_n) \| \leq \sqrt{\lambda^{-1} \varepsilon}, \quad \limsup_n \| x_n^* - y_n^* \| \leq \sqrt{\lambda \varepsilon},$$

and without loss of generality we can replace $\limsup_n$ by $\lim_n$. \hfill \blacksquare

**Open problem:** We don’t know whether the converse (5) $\Rightarrow$ (4) is true.
Combining the last two propositions gives:

**Theorem** Let $X$ Banach and $f : X \to ]-\infty, +\infty]$ lsc, prox-bounded with threshold $\lambda_f \geq 0$. Then:

$$\forall \lambda > \lambda_f, \forall \varepsilon \geq 0, \quad (\partial f)\varepsilon \subset \text{cl} \left( \partial f + \sqrt{\lambda^{-1}\varepsilon}B_X \times \sqrt{\lambda\varepsilon}B_{X^*} \right).$$

Equivalently: for all $\lambda > \lambda_f$ and $\varepsilon \geq 0$, $(x^*, x) \in (\partial f)\varepsilon \Rightarrow 
\exists (x_n^*, x_n) \subset \partial f : \lim_n \|x - x_n\| \leq \sqrt{\lambda^{-1}\varepsilon} \& \lim_n \|x^* - x_n^*\| \leq \sqrt{\lambda\varepsilon}.

In case $\lambda_f = 0$ (in particular for a convex $f$), the wide monotone absorption property is equivalent to the so-called *maximal monotonicity of Brøndsted-Rockafellar type* studied in Simons (1999, 2008) and others, hence the above theorem extends known results for convex functions to the class of prox-bounded non necessarily convex functions, with a more direct proof.
REFERENCES


