BULL. AUSTRAL. MATH. SOC. 90C25, 52A41
VOL. 47 (1993) [333–340]

A FORMULA ON THE SUBDIFFERENTIAL OF THE DECONVOLUTION OF CONVEX FUNCTIONS

M. VOLLE

It is known that, under suitable assumptions, the subdifferential $\partial(f \square g)$ of the infimal convolution of two convex functions f and g coincides with the parallel sum of ∂f and ∂g. We prove that a similar formula holds for the subdifferential of the deconvolution of two convex functions: under appropriate hypothesis it coincides with the parallel star-difference of the sub-differentials of the functions.

1. PRELIMINARIES

In what follows (X, Y) is a couple of locally convex real topological spaces paired in separating duality by a bilinear form $\langle .,. \rangle$, and $\Gamma_0(X)$(respectively $\Gamma_0(Y)$) is the class of convex, lower-semicontinuous proper functions defined on X(respectively Y) with values in $\mathbb{R} \cup \{+\infty\}$. Given $h, k : X \to \overline{\mathbb{R}}$, the inf-convolution $h \square k$ is defined by

$$(h \square k)(x) = \inf_{u \in X} \{h(x-u) \dot{+} k(u)\} \qquad \text{for all } x \in X ,$$

where $\dot{+}$ is the upper extension of the addition to $\overline{\mathbb{R}}$ (that is, $(+\infty)\dot{+}(-\infty) = (+\infty)\dot{-}(+\infty) = +\infty$, see [10]).

The deconvolution, denoted by the symbol $\boxminus$, is a kind of inverse operation for the inf-convolution. It was introduced by Hiriart-Urruty and Mazure [5] in order to solve the inf-convolution equation

$$\text{(1)} \qquad \text{find} \quad \xi \in \overline{\mathbb{R}}^X \quad \text{such that} \quad k \square \xi = h .$$

It is known [9, Corollary 2] that a solution to (1) exists if and only if the function

$$\text{(2)} \qquad x \longmapsto (h \boxminus k)(x) = \sup_{u \in X} \{h(x+u) \underset{\cdot}{-} k(u)\}$$

is one of them. The function defined by (2) is referred to as the deconvolution of h and k. Here the symbol $\underset{\cdot}{-}$ denotes the lower extension of the subtraction to $\overline{\mathbb{R}}$ (that is, $(+\infty)\underset{\cdot}{-}(+\infty) = (+\infty)\underset{\cdot}{+}(-\infty) = -\infty$, [10]).

Received 27th April, 1992

https://doi.org/10.1017/S0004972700012569 Published online by Cambridge University Press

The operation $\boxminus$ has many interesting applications. For instance, taking the deconvolution of convex quadratic forms yields a variational formulation of the parallel subtraction of matrices and operators (for example [8, 13]).

The deconvolution operation is strongly linked to the star-difference of sets. Recall that the star-difference of two subsets A and B of a linear space E is defined by

$$A\overset{*}{-}B = \{x \in E : x + B \subset A\} .$$

By setting
$$E(f) = \{(x,r) \in X \times \mathbb{R} : f(x) \leqslant r\}$$

for the epigraph of $f \in \overline{\mathbb{R}}^E$ and I_C for the indicator function of $C \subset X$ ($I_C(x) = 0$ if $x \in C$, $I_C(x) = +\infty$ if $x \in E \setminus C$), we have then [15, Proposition 6]

(3) $$E(h \boxminus k) = E(h)\overset{*}{-}E(k)$$

(4) $$I_A \boxminus I_B = I_{A\overset{*}{-}B} \qquad \text{if} \quad B \neq \emptyset .$$

In the context of epigraphical analysis [1], formula (3) suggests another terminology for the deconvolution operation, namely the epigraphical difference or, better, the epigraphical star-difference.

In connection with Fenchel's duality theory, the deconvolution operation enjoys some noteworthy properties. Recall that the Fenchel conjugate of $f \in \overline{\mathbb{R}}^X$ is defined by

$$f^*(y) = \sup_{x\in X}\{\langle x,y\rangle - f(x)\} \quad \text{for all} \ \ y \in Y .$$

In a similar way one defines the conjugate of a function in $\overline{\mathbb{R}}^Y$. A fundamental result concerning the conjugacy operation is that

$$f = f^{**} \quad \text{for all} \ \ f \in \Gamma_0(X) .$$

Now, according to Hiriart-Urruty [4, Theorem 2.2], the formula

(5) $$\left(h^* \dot{-} k^*\right)^* = h \boxminus k$$

holds for all $h, k \in \Gamma_0(X)$. It follows that

(6) $$(h \boxminus k)^* = \left(h^* \dot{-} k^*\right)^{**} .$$

If $h^* \dot{-} k^*$ turns out to be in $\Gamma_0(Y)$, one can also write

(7) $$(h \boxminus k)^* = h^* \dot{-} k^* .$$

https://doi.org/10.1017/S0004972700012569 Published online by Cambridge University Press

We end this section by recalling some facts and by introducing some notation. For any extended-real-valued function $\xi \in \overline{\mathbb{R}}^Y$, the ε-subdifferential ($\varepsilon \geqslant 0$) of ξ at $y \in \xi^{-1}(\mathbb{R})$ is the set

$$\partial_\varepsilon \xi(y) = \{x \in X : \forall v \in Y : \xi(v) - \xi(y) \geqslant \langle x, v - y\rangle - \varepsilon\} .$$

For $\varepsilon = 0$ we set, as usual, $\partial_0 \xi(y) = \partial \xi(y)$. In connection with this concept, the following classical property will be used later on (see for example [**6**, p.351]):

LEMMA 1. *For any $\xi \in \overline{\mathbb{R}}^Y$ and $y \in Y$ such that $\xi(y) = \xi^{**}(y) \in \mathbb{R}$ one has*

$$\partial \xi(y) = \partial \xi^{**}(y) .$$

PROOF: As we always have $\xi^{**} \leqslant \xi$, the inclusion $\supset$ is obvious. Now, for any $x \in \partial \xi(y)$, the affine continuous form $\langle x, .\rangle - \langle x, y\rangle + \xi^{**}(y)$ is smaller than ξ. As ξ^{**} coincides with the upper hull of all affine continuous minorants of ξ, we have $\langle x, .\rangle - \langle x, y\rangle + \xi^{**}(y) \leqslant \xi^{**}$, that is to say, $x \in \partial \xi^{**}(y)$. ∎

The directional derivative of a function $\xi \in \overline{\mathbb{R}}^Y$ at a point $y \in \xi^{-1}(\mathbb{R})$ is defined, when it exists, by

$$\xi'(y, d) = \lim_{t \to 0_+} t^{-1}(\xi(y + td) - \xi(y)) \quad \text{for all } d \in Y ;$$

the lower subdifferential of ξ at y is the set (for example [**14**])

$$\partial^- \xi(y) = \{x \in X : \langle x, .\rangle \leqslant \xi'(y, .)\} .$$

The above set obviously contains $\partial \xi(y)$ and coincides with $\partial \xi(y)$ when ξ is convex; in that case (ξ convex) it is well known (for example [**6**, p.354]) that $\xi'(y, .)$ is a sublinear function whose Fenchel conjugate is the indicator function of $\partial \xi(y)$; in other words

(8) $$\left(\xi'(y, .)\right)^* = I_{\partial \xi(y)} .$$

If, moreover, ξ is continuous at y, then $\xi'(y, .)$ is finitely valued, continuous, and one has

(9) $$\xi'(y, .) = \left(I_{\partial \xi(y)}\right)^* .$$

2. On the subdifferential of the difference of two convex functions

Let φ and ψ in $\overline{\mathbb{R}}^Y$ be two convex functions, finite at the point $y \in Y$. In connection with the subdifferentiability of the difference $\varphi - \psi$, there are two formulas worth mentioning:

(10) $$\partial(\varphi \dot{-} \psi)(y) = \bigcap_{\varepsilon > 0} \partial_\varepsilon \varphi(y) \overset{*}{-} \partial_\varepsilon \psi(y)$$

(11) $$\partial^-(\varphi - \psi)(y) = \partial \varphi(y) \overset{*}{-} \partial \psi(y) .$$

The first one is due to Martinez-Legaz and Seeger [**7**, Theorem 1] and applies to arbitrary functions φ and ψ in $\Gamma_0(Y)$; the second one has been mentioned by Ellaia [**3**, p.94] for φ and ψ convex on $\mathbb{R}^n$ and finitely valued. The next lemma extends the second formula to our general setting.

Lemma 2. *Let $\varphi, \psi \in \overline{\mathbb{R}}^Y$ be convex functions, finite at $y \in Y$, and assume that ψ is continuous at y. Then*

$$\partial^-(\varphi \dot{-} \psi)(y) = \partial^-\Big(\varphi \underset{\cdot}{-} \psi\Big)(y) = \partial\varphi(y) \overset{*}{-} \partial\psi(y) .$$

Proof: Let us consider only the case $\dot{-}$. Each of the following lines is equivalent to $x \in \partial\varphi(y) \overset{*}{-} \partial\psi(y)$:

$$\forall u \in \partial\psi(y) : x + u \in \partial\varphi(y)$$
$$\forall u \in \partial\psi(y) : \langle x + u, .\rangle \leqslant \varphi'(y,.) \qquad \big(\text{as } \partial\varphi(y) = \partial^-\varphi(y)\big)$$
$$\langle x, .\rangle + \big(I_{\partial\psi(y)}\big)^* \leqslant \varphi'(y,.) \qquad (\text{by taking the supremum for } u \in \partial\psi(y))$$
$$\langle x, .\rangle + \psi'(y,.) \leqslant \varphi'(y,.) \qquad (\text{from } (9))$$
$$\langle x, .\rangle \leqslant \varphi'(y,.) - \psi'(y,.) \qquad \Big(\text{ as } \psi'(y,.) \text{ is finitely valued}\Big)$$
$$\langle x, .\rangle \leqslant (\varphi \dot{-} \psi)'(y,.)$$
$$x \in \partial^-(\varphi \dot{-} \psi)(y) .$$

∎

Before passing to next section we record here a by-product of this lemma.

Corollary. *Let φ, ψ be in $\Gamma_0(Y)$; assume that $\varphi \dot{-} \psi$ is convex, finite at y, and ψ is continuous at y; then*

$$\bigcap_{\varepsilon > 0} \partial_\varepsilon\varphi(y) \overset{*}{-} \partial_\varepsilon\psi(y) = \partial\varphi(y) \overset{*}{-} \partial\psi(y) .$$

Proof: As $\partial^-(\varphi \dot{-} \psi) = \partial(\varphi \dot{-} \psi)$, it suffices to apply formula (10) and Lemma 2. ∎

3. On the subdifferential of the deconvolution

Let us present now the main results of this note. Given $h, k \in \Gamma_0(X)$, the parallel sum of ∂h and ∂k is defined by (see for example [**12**])

$$\partial h \,\square\, \partial k = \Big((\partial h)^{-1} + (\partial k)^{-1}\Big)^{-1} .$$

https://doi.org/10.1017/S0004972700012569 Published online by Cambridge University Press

Here $(\partial h)^{-1}$ is set for the inverse of the multivalued operator ∂h; in other words:

$$y \in (\partial h \,\square\, \partial k)(x) \Leftrightarrow x \in (\partial h)^{-1}(y) + (\partial k)^{-1}(y)$$

where + denotes the Minkowski vectorial addition. It turns out that, under appropriate constraint qualifications [**11**, **12**], the formula $\partial(h \,\square\, k) = \partial h \,\square\, \partial k$ holds. It is tempting to ask wether or not there is a similar formula for the deconvolution operator. To this end, let us introduce the notion of parallel star difference for subdifferentials.

DEFINITION: Let h and k be in $\Gamma_0(X)$. The parallel star difference of the subdifferentials ∂h and ∂k is the multivalued operator $\partial h \boxminus \partial k$ defined by

$$\partial h \boxminus \partial k = \left((\partial h)^{-1} \overset{*}{-} (\partial k)^{-1}\right)^{-1},$$

that is, for any $(x,y) \in X \times Y$,

$$y \in (\partial h \boxminus \partial k)(x) \Leftrightarrow x \in (\partial h)^{-1}(y) \overset{*}{-} (\partial k)^{-1}(y) .$$

In [**5**, Proposition 7] one finds a lower estimate for the subdifferential of two finitely valued convex functions f, g on $\mathbb{R}^n$; namely it is shown that

$$\partial(g \boxminus f)(x) \;\supset \bigcup_{(x_1,x_2)\in Ax} \partial g(x_1) \cap \partial f(x_2) ,$$

where $Ax = \{(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n : x = x_1 - x_2,\ (g \boxminus f)(x) = g(x_1) - f(x_2)\}$. As pointed out to me by A. Seeger, the condition $(x_1, x_2) \in Ax$ yields the inclusion $\partial g(x_1) \subset \partial f(x_2)$. Moreover, the convexity assumption on f and g is superfluous:

PROPOSITION. *Let f, g be arbitrary extended real valued functions on X. Then, for any $x \in X$, we have*

$$\partial(g \boxminus f)(x) \;\supset \bigcup_{v\in E(x)} \partial g(v) ,$$

where $E(x) = \{v \in X : g(v) - f(v-x) = (g \boxminus f)(x) \in \mathbb{R}\}$.

PROOF: Let v be in $E(x)$ and $y \in \partial g(v)$; then $g(v)$ and $f(v-x)$ are real numbers and we have,

$$(g \boxminus f)(z) \geqslant g(z+v-x) - f(v-x) \quad \text{for all } z \in X .$$

Hence,

$$(g \boxminus f)(z) - (g \boxminus f)(x) \geqslant g(z+v-x) - f(v-x) - g(v) + f(v-x) \quad \text{for all } z \in X ,$$

and finally

$$(g \boxminus f)(z) - (g \boxminus f)(x) \geqslant g(z+v-x) - g(v) \geqslant \langle z-x, y\rangle \quad \text{for all } z \in X .$$

This shows that $y \in \partial(g \boxminus f)(x)$. $\square$

The next result provides an upper estimate for the subdifferential of the deconvolution of two convex functions $h, k \in \Gamma_0(X)$ in terms of the parallel star difference of ∂h and ∂k; it involves the set

$$C(h,k) = \{y \in Y : k^* \text{ is finite and continuous at } y, \text{ and } (h^* \dot{-} k^*)(y) = (h^* \dot{-} k^*)^{**}(y)\}$$

Theorem 1. *Let X, Y be locally convex spaces in separating duality, and let $h, k \in \Gamma_0(X)$. Then, for all $x \in X$, we have*

$$\partial(h \boxminus k)(x) \cap C(h,k) \subset (\partial h \boxminus \partial k)(x) .$$

Proof: Assume that $y \in \partial(h \boxminus k)(x) \cap C(h,k)$. We have to show that $x \in (\partial h)^{-1}(y) \overset{*}{-} (\partial k)^{-1}(y)$. As $y \in \partial(h \boxminus k)(x)$ we have $x \in \partial(h \boxminus k)^*(y)$. Now, from (6), $(h \boxminus k)^* = (h^* \dot{-} k^*)^{**}$; then $x \in \partial(h^* \dot{-} k^*)^{**}(y)$; as $y \in C(h,k)$ it follows from Lemma 1 that $x \in \partial(h^* \dot{-} k^*)(y)$ and, a fortiori, $x \in \partial^-(h^* \dot{-} k^*)(y)$. So, by Lemma 2, we obtain $x \in \partial h^*(y) \overset{*}{-} \partial k^*(y)$. $\square$

Let us give an example showing that Theorem 1 cannot be improved without additional assumptions. Take for X an Hilbert space with closed unit ball B, $h = \| \, \|$, $k = \left(\| \, \|^2\right)/2$. We have then by (6) $(h \boxminus k)^* = \left(I_B - \left(\| \, \|^2\right)/2\right)^{**} = I_B - 1/2$ so that $h \boxminus k = \| \, \| + 1/2$. Note also that $C(h,k) = \{y \in X : \|y\| \geqslant 1\}$. For the subdifferentials we have, on one hand,

$$\partial(h \boxminus k)(x) = \begin{cases} \dfrac{x}{\|x\|} & \text{if } \quad x \neq 0 \\ B & \text{if } \quad x = 0 \end{cases}$$

and, on the other hand, $y \in (\partial h \boxminus \partial k)(x)$ if and only if

$$x \in \partial I_B(y) \overset{*}{-} \partial\left(\frac{\| \, \|^2}{2}\right)(y) = \partial I_B(y) - y = \begin{cases} \emptyset & \text{if } \quad \|y\| > 1 \\ [-1, +\infty[y & \text{if } \quad \|y\| = 1 \\ -y & \text{if } \quad \|y\| < 1 \end{cases} .$$

In particular,

$$\text{for } \quad \|x\| = 1 \qquad \partial(h \boxminus k)(x) = \{x\} \underset{\neq}{\subset} (\partial h \boxminus \partial k)(x) = \{x, -x\}$$

$$\text{for } \quad x = 0 \qquad \partial(h \boxminus k)(0) = B \underset{\neq}{\supset} (\partial h \boxminus \partial k)(0) = \{y \, : \, \|y\| = 1\} \cup \{0\} .$$

With stronger assumptions it is possible to give an exact formula for $\partial(h \boxminus k)$:

https://doi.org/10.1017/S0004972700012569 Published online by Cambridge University Press

Theorem 2. *Let X, Y be locally convex spaces paired in separating duality, and let $h, k \in \Gamma_0(X)$. Assume that k^* is finite and continuous over Y and that $h^* - k^*$ is convex. Then,*

$$\partial(h \boxminus k) = \partial h \boxminus \partial k .$$

Proof: Since for each $f \in \Gamma_0(X)$ one has $(\partial f)^{-1} = \partial f^*$, we easily obtain the equivalence between the assertions below:

$$y \in \partial(h \boxminus k)(x)$$

$$x \in \partial(h \boxminus k)^*(y)$$

$$x \in \partial(h^* - k^*)(y) \quad \text{(by (5) as } h^* - k^* \text{ is convex proper lower semicontinuous)}$$

$$x \in \partial h^*(y) \overset{*}{-} \partial k^*(y) \quad \text{(from Lemma 2)}$$

$$x \in (\partial h)^{-1}(y) \overset{*}{-} (\partial k)^{-1}(y) .$$

□

Example: Let us take for X a Hilbert space, $h \in \Gamma_0(X)$, $k \in \Gamma_0(X)$. Assume that k^* is finite over X (hence continuous) and suppose that $h^* - k^*$ is strongly convex: there exists $t > 0$ and $f \in \Gamma_0(X)$ such that $h^* - k^* = f^* + \left(t \|\ \|^2\right)/2$. We have then by (5)

$$h \boxminus k = \left(f^* + \frac{t\|\ \|^2}{2}\right)^* = f \,\Box\, \frac{\|\ \|^2}{2t} .$$

So, $h \boxminus k$ coincides with the Moreau-Yosida regularisation of $f \in \Gamma_0(X)$ (for example [**2**, p.195]). It follows that $h \,\Box\, k$ is continuously differentiable. As h^* is also strongly convex, h is continuously differentiable and we have, applying Theorem 2,

$$\nabla(h \boxminus k) = \nabla h \boxminus \partial k .$$

References

[1] H. Attouch and R.J.B. Wets, 'Epigraphical analysis', in *Analyse non linaire*, (H. Attouch, J.-P. Aubin, F.H. Clarke, I. Ekeland, Editors) (Gauthier-Villars, Paris, 1989), pp. 73–100.

[2] J.-P. Aubin and I. Ekeland, *Applied nonlinear analysis* (Wiley, New York, 1984).

[3] R. Ellaia, *Contribution à l'analyse et l'optimisation de difference de fonctions convexes*, Thèse (Université Paul Sabatier, Toulouse, 1984).

[4] J.-B. Hiriart-Urruty, 'A general formula on the conjugate of the difference of functions', *Canad. Math. Bull.* **29** (1986), 482–485.

[5] J.-B. Hiriart-Urruty and M.-L. Mazure, 'Formulations variationnelles de l'addition parallèle et de la soustraction parallèle d'opérateurs semi-définis positifs', *C.R. Acad. Sci. Paris Série I* **302** (1986), 527–530.

[6] P.-J. Laurent, *Approximation et optimisation* (Hermann, Paris, 1972).

[7] J.E. Martinez-Legaz and A. Seeger A formula on the approximate subdifferential of the difference of convex functions, *Bull. Austral. Math. Soc.* **45** (1992), 37–41.

[8] M.-L. Mazure, 'La soustraction parallèle d'opérateurs interpretée comme déconvolution de formes quadratiques convexes', *Optimization* **18** (1987), 465–484.

[9] M.-L. Mazure and M. Volle, 'Equations inf-convolutives et conjugaison de Moreau-Fenchel', *Ann. Fac. Sci. Toulouse Math* **12** (1991), 103–126.

[10] J.-J. Moreau, 'Inf-convolution, sous-additivité, convexité des fonctions numériques', *J. Math. Pures Appl.* **49** (1970), 109–154.

[11] A. Moudafi, *Convolution des opérateurs monotones et convergence variationnelle* (Université de Clermont-Ferrand II, 1991). (Preprint).

[12] G.B. Passty, 'The parallel sum of nonlinear monotone operators', *Nonlinear Anal.* **10** (1986), 215–227.

[13] E.L. Pekarev and L. Smul'jan, 'Parallel addition and parallel subtraction of operators', *Math. USSR-Izv.* **10** (1976), 351–370.

[14] J.-P. Penot, 'Calcul sous-différentiel et optimisation', *J. Funct. Anal.* **27** (1978), 248–276.

[15] M. Volle, 'Concave duality : application to problems dealing with difference of functions', *Math. Programming* **41** (1988), 261–278.

University of Avignon
33 rue L. Pasteur
84000 Avignon
France