

Conciliating Generalized Derivatives

Jean-Paul Penot

Abstract As a new illustration of the versatility of abstract subdifferentials we examine their introduction in the field of analysis of second-order generalized derivatives. We also consider some calculus rules for some of the various notions of such derivatives and we give an account of the effect of the Moreau regularization process.

Keywords Calculus rules • Coderivatives • Epi-derivatives • Generalized derivatives • Subdifferentials • Variational analysis

1 Introduction

It is likely that many mathematicians and users have been disoriented by the abundance of concepts in nonsmooth analysis. The following quotation of F.H. Clarke's review of the book [45] attests that: "In recent years, the subject has known a period of intense abstract development which has led to a rather bewildering array of competing and unclearly related theories." Even in leaving apart some important approaches such as Demyanov's quasidifferentials [7, 8], Jeyakumar and Luc's generalized Jacobians [20], Warga's derivate containers [46] and in focussing on subdifferentials, one is faced with very different constructs. It may be useful to be aware of their specific features in order to choose the concept adapted to the problem one has to solve. That is the point of view adopted in the author's book [38]. It appears there that even if the approaches are different, a number of important properties are shared by all usual subdifferentials, at least on suitable spaces.

Thus, several researchers endeavored to give a synthetic approach to generalized derivatives in proposing a list of properties that can be taken as axioms allowing

J.-P. Penot (✉)

Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie (Paris 6), Paris, France
e-mail: penot@ann.jussieu.fr

to ignore the various constructions. Such a list may be more or less complete depending on the needs or the aims. Among the many attempts in this direction, we quote the recent papers [18, 41] as they gather the most complete lists to date; see also their references.

Among the results using a general subdifferential ∂ , we note mean value theorems [36], characterizations of convexity [5, 6], or generalized convexity [33, 35, 37, 42], optimality conditions using Chaney type second derivatives [40]. It is the purpose of the present paper to show that a general approach using an arbitrary subdifferential can also be workable for general second-order derivatives.

It is amazing that second-order derivatives can be considered for functions that are not even differentiable. In fact, many approaches can be adopted and it is a second aim of the present paper to relate such derivatives. Since the methods are so diverse (primal, dual, using second-order expansions or not), such an aim is not obvious. In fact, several results are already known; we leave them apart. For instance, for the links between parabolic second derivatives with epi-derivatives, we refer to [13–15, 30, 39] and [45, Sect. 13.J]. We do not consider generalized Hessians and generalized Jacobians of the derivative that have been the object of a strong attention during the last decades (see [29] and its references). We also leave apart second-order derivatives involving perturbations of the nominal point as in [4, 25], [45, Proposition 13.56], [47], and second-order derivatives aimed at convex functions [10–12].

2 Calculus of Second-Order Epi-derivatives

The (lower) *second-order epi-derivative* of $f : X \rightarrow \overline{\mathbb{R}}$ at $x \in f^{-1}(\mathbb{R})$ relative to $x^* \in X^*$ has been introduced by Rockafellar in [43]; it is given by

$$\forall u \in X \quad f''(x, x^*, u) := \liminf_{(t, tu') \rightarrow (0+, u)} \frac{2}{t^2} (f(x + tu') - f(x) - \langle x^*, tu' \rangle).$$

We also write $f''_{x, x^*}(u) = f''(x, x^*, u)$. When the directional derivative $f'(x)$ of f at x exists, we write f''_x instead of f''_{x, x^*} with $x^* := f'(x)$. Obviously, if for all $u \in X$ one has $f''(x, x^*, u) > -\infty$, one has $x^* \in \partial_D f(x)$, the directional subdifferential of f at x . Moreover, if f is twice directionally differentiable at x in the sense that there exist some $x^* := f'(x) \in X^*$ and some $D^2 f(x) \in L^2(X, X; \mathbb{R})$ such that for all $u \in X$ one has

$$f'(x)(u) = \lim_{(t, tu') \rightarrow (0+, u)} \frac{f(x + tu') - f(x)}{t},$$

$$D^2 f(x)(u, u) = \lim_{(t, tu') \rightarrow (0+, u)} \frac{f(x + tu') - f(x) - \langle x^*, tu' \rangle}{t^2/2},$$

then $f''(x, x^*, u) = D^2f(x)(u, u)$. More generally, let us say that a continuous linear map $H : X \rightarrow X^*$ is a *directional subhessian* of f at $(x, x^*) \in X \times X^*$ if there exists a function $r : \mathbb{R} \times X \rightarrow \mathbb{R}$ such that $\lim_{(t, u') \rightarrow (0+, u)} t^{-2}r(t, u') = 0$ for all $u \in X$ and

$$\forall u' \in X, \forall t \in \mathbb{P} \quad f(x + tu') \geq f(x) + \langle x^*, tu' \rangle + t^2 \langle Hu', u' \rangle + r(t, u').$$

If there exists a function $s : X \rightarrow \mathbb{R}$ such that $r(t, u) = s(tu)$ for all $(t, u) \in \mathbb{R} \times X$ and $\|x\|^{-2}s(x) \rightarrow 0$ as $x \rightarrow 0$ in $X \setminus \{0\}$ we say that H is a *firm subhessian* or a Fréchet subhessian. Then one sees that H is a directional subhessian of f at (x, x^*) if, and only if, for all $u \in X$ one has

$$f''(x, x^*, u) \geq \langle Hu, u \rangle.$$

We refer to [32] for the links between subhessians and conjugacy and to [19] for calculus rules for (limiting) subhessians.

Second-order epi-derivatives are useful to get second-order optimality conditions (see [43–45]).

The calculus rules for such derivatives are not simple (see [16, 17, 27, 31, 34, 45]). However, some estimates can be easily obtained and have some usefulness. Clearly, if $f = g + h$ and if $x^* = y^* + z^*$, one has

$$f''(x, x^*, u) \geq g''(x, y^*, u) + h''(x, z^*, u)$$

for all $u \in \text{dom}g''_{x, y^*} \cap \text{dom}h''_{x, z^*}$. Suppose $f := h \circ g$, where $g : X \rightarrow Y$ is twice directionally differentiable at x and $h : Y \rightarrow \overline{\mathbb{R}}$ is finite at $y := g(x)$, given $y^* \in Y^*$ and $x^* := y^* \circ g'(x)$, one has

$$f''(x, x^*, u) \geq h''(y, y^*, g'(x)u) + \langle y^*, D^2g(x)(u, u) \rangle,$$

as shows the fact that

$$w(t, u') := \frac{2}{t^2}(g(x + tu') - g(x) - tg'(x)u' - \frac{t^2}{2}D^2g(x)(u, u)) \xrightarrow{(t, u') \rightarrow (0+, u)} 0,$$

so that $g'(x)u' + (1/2)t(D^2g(x)(u, u) + w(t, u')) \rightarrow g'(x)u$ as $(t, u') \rightarrow (0+, u)$.

Let us consider the relationships of such derivatives with graphical derivatives of the subdifferential. Recall that the (outer) *graphical derivative* at $(x, y) \in \text{gph}(F)$ of a multimap $F : X \rightrightarrows Y$ between two normed spaces is the multimap $DF(x, y)$ whose graph is the (weak) tangent cone at (x, y) to the graph of F . In other terms,

$$\forall u \in X \quad DF(x, y)(u) := \text{seq-weak-} \limsup_{(t, u') \rightarrow (0+, u)} \Delta_t F(x, y)(u'),$$

where

$$\Delta_t F(x, y)(u') := \frac{1}{t}(F(x + tu') - y).$$

We say that F has a *graphical derivative* at (x, y) or that F is (weakly) proto-differentiable at (x, y) if the sequential weak limsup of $\Delta_t F(x, y)$ as $t \rightarrow 0_+$ coincides with $\liminf_{t \rightarrow 0_+} \Delta_t F(x, y)$.

For $f : X \rightarrow \overline{\mathbb{R}}$ finite at x and $x^* \in \partial f(x)$, let us introduce the differential quotient

$$\forall u \in X \quad \Delta_t^2 f_{x,x^*}(u) := \frac{2}{t^2} [f(x+tu) - f(x) - \langle x^*, tu \rangle].$$

We say that $f : X \rightarrow \overline{\mathbb{R}}$ finite at x has a *second-order epi-derivative* at (x, x^*) with $x^* \in \partial f(x)$ (in the sense of Mosco) if the sequential weak limsup as $t \rightarrow 0_+$ of the epigraph of $\Delta_t^2 f_{x,x^*}$ coincides with the liminf as $t \rightarrow 0_+$ of the epigraph of $\Delta_t^2 f_{x,x^*}$.

The function $f : X \rightarrow \overline{\mathbb{R}}$ is said to be *paraconvex* around x if it is lower semicontinuous, finite at x and if there exist $c > 0, r > 0$ such that the function $f + (c/2) \|\cdot\|^2$ is convex on the ball $B(x, r)$ with center x and radius r .

Assertion (a) of the next proposition completes [45, Lemma 13.39]. Assertion (b) extends [45, Theorem 13.40] to the infinite dimensional case and [9] to a nonconvex case. Our assumptions are slightly different from the assumptions of [23] dealing with primal lower-nice functions, a class that is larger than the class of paraconvex functions but requires a more complex analysis. See also [3, 21–23, 28].

Proposition 1.

(a) For any function $f : X \rightarrow \overline{\mathbb{R}}$ on a normed space X that is finite at $x \in X$ and for all $x^* \in \partial f(x)$ one has

$$D(\partial f)(x, x^*)(u) = \text{seq-weak-} \limsup_{(t, u') \rightarrow (0_+, u)} \partial \left(\frac{1}{2} \Delta_t^2 f_{x,x^*} \right) (u'). \quad (1)$$

(b) If X is a Hilbert space and if f is paraconvex around x , then ∂f has a graphical derivative $D(\partial f)(x, x^*)$ at (x, x^*) if, and only if, f has a second-order epi-derivative f''_{x,x^*} at (x, x^*) . Then

$$D(\partial f)(x, x^*)(u) = \frac{1}{2} \partial f''_{x,x^*}(u).$$

Let us observe that when f has a second-order epi-derivative at (x, x^*) that is quadratic and continuous this relation ensures that $D(\partial f)(x, x^*)$ is the linear and continuous map from X into X^* corresponding to this quadratic form.

Proof.

(a) We first observe that the calculus rules of subdifferentials ensure that

$$\partial \left(\frac{1}{2} \Delta_t^2 f_{x,x^*} \right) (\cdot) = \Delta_t(\partial f)(x, x^*)(\cdot).$$

Passing to the (sequential weak) limsup as $t \rightarrow 0_+$ yields equality on the graphs and relation (1).

- (b) Changing f into $g := f + (c/2) \|\cdot\|^2$ for some appropriate $c > 0$ and observing that $(c/2) \|\cdot\|^2$ is twice continuously differentiable, we reduce the proof to the case f is convex. Now we note that for all $t > 0$ we have $(0, 0) \in \partial(\frac{1}{2}\Delta_r^2 f_{x,x^*})$ and that $(\Delta_r^2 f_{x,x^*})(0) = 0$. Thus we can apply the equivalence of Attouch's theorem ([1, Theorem 3.66]) and get the conclusion. \square

3 Epi-derivatives and Regularization

Let us consider the interplay between second-order epi-differentiation and the Moreau regularization (or Moreau envelope). The *Moreau envelope* $r f := e_r f$ of a function $f : X \rightarrow \mathbb{R}_\infty := \mathbb{R} \cup \{\infty\}$ on a Hilbert space X given by

$$(e_r f)(x) = \inf_{w \in X} \left(f(w) + \frac{1}{2r} \|x - w\|^2 \right).$$

It is well known (see [2, Proposition 12.29], [38, Theorem 4.124] for instance) that if f is paraconvex around $x \in X$ and *quadratically minorized* in the sense that for some $a \in \mathbb{R}, b \in \mathbb{R}$ one has $f \geq b - a \|\cdot\|^2$, then there exists some $\bar{r} > 0$ such that for all $r \in]0, \bar{r}[$ the Moreau envelope $e_r f$ is differentiable at x with gradient $\nabla(e_r f)(x) = x_r := (1/r)(x - p_r)$ where $p_r := p_r(x) \in X$ is such that $(e_r f)(x) = f(p_r) + (1/2r) \|x - p_r\|^2$. The following result has been obtained in [9, Theorem 4.3] for a lower semicontinuous, proper, convex function f under the additional assumption that f is twice epi-differentiable at p_r . The method of [9, Theorem 4.3] uses duality and in particular Attouch's theorem, whereas we take a direct approach. We denote by $J : X \rightarrow X^*$ the Riesz isomorphism characterized by $\langle J(x), w \rangle = \langle x | w \rangle$, where $\langle \cdot | \cdot \rangle$ is the scalar product of X .

Proposition 2. *Let X be a Hilbert space and let $f : X \rightarrow \overline{\mathbb{R}}$ be paraconvex around $x \in X$ and quadratically minorized. Then, there exists some $\bar{r} > 0$ such that for all $r \in]0, \bar{r}[$ the Moreau envelope $e_r f$ is differentiable at x with gradient $\nabla(e_r f)(x) = x_r := (1/r)(x - p_r)$ where $p_r := p_r(x) \in X$ is such that $(e_r f)(x) = f(p_r) + (1/2r) \|x - p_r\|^2$. Moreover, the sequential weak (lower) second derivative of the Moreau envelope of f is related to the Moreau envelope of the (lower) sequential weak second derivative of f by the following relation in which $x_r^* := J(x_r) \in \partial f(p_r)$ is the derivative of $e_r f$ at x :*

$$\forall u \in X \quad \frac{1}{2}(e_r f)''(x, x_r^*, u) = e_r \left(\frac{1}{2} f''(p_r, x_r^*, \cdot) \right) (u).$$

In particular, when f is lower semicontinuous, proper, convex, and twice epi-differentiable at p_r for x_r^* , the function $(e_r f)''(x, x_r^*, \cdot)$ is continuous (and even of class C^1): in such a case $(1/2)f''(p_r, x_r^*, \cdot)$ is lower semicontinuous, proper, convex, so that $e_r((1/2)f''(p_r, x_r^*, \cdot))$ is of class C^1 . For the proof we need a result of independent interest about the interchange of minimization with a sequential weak lower epi-limit.

Lemma 1. *Let M be a metric space, let $S \subset M$, $\bar{s} \in \text{cl}(S)$ and let Z be a reflexive Banach space endowed with its weak topology. Let $g : S \times Z \rightarrow \bar{\mathbb{R}}$ be coercive in its second variable, uniformly for $s \in S \cap V$, where V is a neighborhood of t in M . Then*

$$\liminf_{s(\in S) \rightarrow \bar{s}} (\inf_{z \in Z} g(s, z)) = \inf_{z \in Z} (\text{seq-} \liminf_{(s, z')(\in S \times Z) \rightarrow (\bar{s}, z)} g(s, z'))$$

Proof. The inequality

$$\liminf_{s(\in S) \rightarrow \bar{s}} (\inf_{z \in Z} g(s, z)) \leq \inf_{z \in Z} (\text{seq-} \liminf_{(s, z')(\in S \times Z) \rightarrow (\bar{s}, z)} g(s, z'))$$

stems from the fact that for all $\bar{z} \in Z$, $(z_n) \rightarrow \bar{z}$ for the weak topology and $(s_n) \rightarrow \bar{s}$ in S one has $\inf_{z \in Z} g(s_n, z) \leq g(s_n, z_n)$ for all $n \in \mathbb{N}$, hence $\liminf_n (\inf_{z \in Z} g(s_n, z)) \leq \liminf_n g(s_n, z_n)$.

Let $r > \liminf_{s(\in S) \rightarrow \bar{s}} (\inf_{z \in Z} g(s, z))$. Given a sequence $(\varepsilon_n) \rightarrow 0_+$ one can find some $s_n \in B(\bar{s}, \varepsilon_n) \cap S$, $z_n \in Z$ such that $r > g(s_n, z_n)$. We may suppose $B(\bar{s}, \varepsilon_n) \subset V$ for all n . The coercivity assumption ensures that (z_n) is bounded. Taking a subsequence if necessary, we may suppose (z_n) has a weak limit z . Then we get

$$r \geq \liminf_n g(s_n, z_n) \geq \text{seq-} \liminf_{(s, z')(\in S \times Z) \rightarrow (\bar{s}, z)} g(s, z')$$

and $r \geq \inf_{z \in Z} (\text{seq-} \liminf_{(s, z')(\in S \times Z) \rightarrow (\bar{s}, z)} g(s, z'))$. □

For $(t, u) \in \mathbb{P} \times X$, assuming the classical fact that ${}^r f = e_r f$ is differentiable at x with derivative x_r^* , we adopt the simplified notation

$$\Delta_r^2 (e_r f)_x (u) := \frac{2}{t^2} ((e_r f)(x + tu) - (e_r f)(x) - \langle x_r^*, tu \rangle).$$

Then we have the following exchange property generalizing [45, Lemma 13.39].

Lemma 2. *With the preceding notation, one has*

$$\frac{1}{2} \Delta_t^2 (e_r f)_x = e_r \left(\frac{1}{2} \Delta_t^2 f_{p_r, x_r^*} \right).$$

Proof. Setting $w = p_r + tz$ with $t \in \mathbb{P}$, $z \in X$, for all $u \in X$ we have $\langle x_r^*, tu \rangle = \frac{t}{r} \langle x - p_r \mid u \rangle$, $(e_r f)(x) = f(p_r) + \frac{1}{2r} \|x - p_r\|^2$,

$$\begin{aligned} (e_r f)(x + tu) &= \inf_{w \in X} \left(f(w) + \frac{1}{2r} \|x + tu - w\|^2 \right) \\ &= \inf_{z \in X} \left(f(p_r + tz) + \frac{1}{2r} \|x - p_r + t(u - z)\|^2 \right), \end{aligned}$$

$$\begin{aligned} & \frac{1}{2r} \|x - p_r + t(u - z)\|^2 - \frac{1}{2r} \|x - p_r\|^2 - \frac{t}{r} \langle x - p_r \mid u \rangle \\ &= -\frac{t}{r} \langle x - p_r \mid z \rangle + \frac{t^2}{2r} \|u - z\|^2 \end{aligned}$$

hence

$$\begin{aligned} & \frac{1}{2} \Delta_r^2 (e_r f)_x (u) \\ &= \frac{1}{t^2} \inf_{w \in X} \left[f(w) - f(p_r) + \frac{1}{2r} \|x + tu - w\|^2 - \frac{1}{2r} \|x - p_r\|^2 - \langle x_r^*, tu \rangle \right] \\ &= \frac{1}{t^2} \inf_{z \in X} \left[f(p_r + tz) - f(p_r) - \frac{t}{r} \langle x - p_r \mid z \rangle + \frac{t^2}{2r} \|u - z\|^2 \right] \\ &= \inf_{z \in X} \left[\frac{1}{t^2} (f(p_r + tz) - f(p_r) - \langle x_r^*, tz \rangle) + \frac{1}{2r} \|u - z\|^2 \right] \\ &= e_r \left(\frac{1}{2} \Delta_t^2 f_{p_r, x_r^*} \right) (u). \end{aligned}$$

Proof of Proposition 2. The first assertion is deduced from the convex case by considering $f + c \|\cdot\|^2$ for some appropriate $c > 0$ (see [2], [45, Theorem 2.26]). For the second one, in Lemma 1 let us set $Z := X$, $S := \mathbb{R} \times X$, $M := \mathbb{P} \times X$, $s := (t, u')$, $\bar{s} := (0, u)$, $g := h + k$ with

$$\begin{aligned} h(s, z) &:= \Delta_t^2 f(p_r, x_r^*, z) := (1/t^2)(f(p_r + tz) - f(p_r) - \langle x_r^*, tz \rangle), \\ k(s, z) &:= (1/2r) \|u' - z\|^2. \end{aligned}$$

It remains to show that for all $z \in Z$ one has

$$\text{seq} - \liminf_{(s, z') \in (S \times Z) \rightarrow (\bar{s}, z)} g(s, z') = \frac{1}{2} f''(p_r, x_r^*, z) + \frac{1}{2r} \|u - z\|^2. \tag{2}$$

For every sequences $(t_n) \rightarrow 0_+$, $(u_n) \rightarrow u$, $(z_n) \rightarrow z$ in the weak topology of $Z = X$, one has

$$\begin{aligned} \liminf_n g((t_n, u_n), z_n) &\geq \liminf_n h((t_n, u_n), z_n) + \liminf_n k((t_n, u_n), z_n) \\ &\geq \frac{1}{2} f''(p_r, x_r^*, z) + \frac{1}{2r} \|u - z\|^2. \end{aligned}$$

Let $(t_n) \rightarrow 0_+$, $(z_n) \rightarrow z$ (in the weak topology) be such that $(\Delta_{t_n}^2 f(p_r, x_r^*, z_n)) \rightarrow \frac{1}{2} f''(p_r, x_r^*, z)$. Setting $u_n := u + z_n - z$, we note that $(u_n) \rightarrow u$ in the weak topology and $\|u_n - z_n\|^2 = \|u - z\|^2$ for all n , so that

$$\begin{aligned} \frac{1}{2}f''(p_r, x_r^*, z) + \frac{1}{2r} \|u - z\|^2 &= \lim_n h((t_n, u_n), z_n) + \lim_n k((t_n, u_n), z_n) \\ &= \lim_n g((t_n, u_n), z_n) \geq \text{seq} - \liminf_{(s, z') \in S \times Z \rightarrow (\bar{s}, z)} g(s, z'). \end{aligned}$$

Since the reverse inequality stems from the preceding estimate, we get relation (2). \square

4 Second-Order Derivatives via Coderivatives

Let us recall that the *coderivative* of a multimap (or set-valued map) $F : U \rightrightarrows V$ between two normed spaces at $(\bar{u}, \bar{v}) \in F$ (identified with its graph $\text{gph}(F)$) is the multimap $D^*F(\bar{u}, \bar{v}) : V^* \rightrightarrows U^*$ defined by

$$D^*F(\bar{u}, \bar{v})(v^*) := \{u^* \in U^* : (u^*, -v^*) \in N(F, (\bar{u}, \bar{v}))\}.$$

If $\varphi : U \rightarrow V$ is a map of class C^1 between two normed spaces, the coderivative of φ at $(\bar{u}, \varphi(\bar{u}))$, denoted by $D^*\varphi(\bar{u})$ rather than $D^*\varphi(\bar{u}, \varphi(\bar{u}))$, is

$$D^*\varphi(\bar{u}) = \varphi'(\bar{u})^T : V^* \rightarrow U^*,$$

a single-valued (linear) map rather than a multimap. When $\varphi = f'$ is the derivative of a function $f : U \rightarrow \mathbb{R}$ (so that $V = U^*$), denoting by $D^2f : U \rightarrow L(U, U^*)$ the derivative of $\varphi := f'$, we obtain that $D^*f'(\bar{u}) = D^2f(\bar{u})^T$ maps U^{**} into U^* and

$$D^*f'(\bar{u})(u^{**}) = u^{**} \circ \varphi'(\bar{u}) = u^{**} \circ D^2f(\bar{u}) \in U^*.$$

Denoting by $A \mapsto A^b$ the isomorphism from $L(U, U^*)$ onto the space $L^2(U, U; \mathbb{R})$ of continuous bilinear forms on U , we see that the restriction of $\partial^2 f(\bar{u}) := D^*f'(\bar{u})$ to U considered as a subspace of U^{**} is $D^2f(\bar{u})^b$:

$$\forall u \in U \quad \partial^2 f(\bar{u})(u) := D^*f'(\bar{u})(u) = D^2f(\bar{u})^b(u, \cdot) \in U^*.$$

Calculus rules for coderivatives are presented in [24, 26]. In [27] calculus rules for coderivatives of subdifferentials are given in the finite dimensional case. Here we look for extensions to the infinite dimensional case.

Since the coderivative of a multimap is defined through the normal cone to its graph, it is natural to expect that calculus rules for second-order derivatives of functions depend on calculus rules for normal cones under images and inverse images. Let us recall such rules (see [38, Proposition 2.108, Theorem 2.111]).

Proposition 3. *Let U, V, W be Banach spaces, and let $j : U \rightarrow V, k : U \rightarrow W$ be maps of class $C^1, E \subset U, H \subset W, \bar{u} \in E, \bar{v} := j(\bar{u}) \in V, \bar{w} := k(\bar{u}) \in W.$*

If $E := k^{-1}(H)$, and if $k'(\bar{u})(U) = W$, then $N(E, \bar{u}) = k'(\bar{u})^T(N(H, \bar{w}))$.
 If $F := j(E)$, then $N(F, \bar{v}) \subset (j'(\bar{u})^T)^{-1}(N(E, \bar{u}))$.
 If $F := j(k^{-1}(H))$, and if $k'(\bar{u})(U) = W$, then, for all $v^* \in N(F, \bar{v})$, there exists a unique $w^* \in N(H, \bar{w})$ such that $j'(\bar{u})^T(v^*) = k'(\bar{u})^T(w^*)$.

For the second assertion it is not necessary to suppose U and V are complete. Moreover, the differentiability assumptions on j and k can be adapted to the specific subdifferentials that are used. But since we wish to adopt a general approach, we ignore such refinements.

Let us consider a composite function $f := h \circ g$, where X, Y are Banach spaces, $g : X_0 \rightarrow Y$ is of class C^2 on some open subset X_0 of X whose derivative at x around $\bar{x} \in X_0$ is surjective and $h : Y \rightarrow \bar{\mathbb{R}}$ is lower semicontinuous and finite around $g(\bar{x})$. For all usual subdifferentials one has the formula

$$\partial f(x) = g'(x)^T(\partial h(g(x))),$$

where $g'(x)^T$ is the transpose of the derivative $g'(x)$ of g at x .

In order to compute $\partial^2 f(\bar{x}, \bar{x}^*) := D^* \partial f(\bar{x}, \bar{x}^*)$, where $\bar{x}^* \in \partial f(\bar{x})$, $\bar{x}^* = g'(\bar{x})^T(\bar{y}^*)$ with $\bar{y}^* \in \partial h(g(\bar{x}))$, let us denote by F (resp. H) the graph of ∂f (resp. ∂h) and set

$$\begin{aligned} E &:= \{(x, y^*) \in X \times Y^* : (g(x), y^*) \in H\}, \\ j &:= I_X \times g'(x)^T : (x, y^*) \mapsto (x, y^* \circ g'(x)), \\ k &:= g \times I_{Y^*} : (x, y^*) \mapsto (g(x), y^*). \end{aligned}$$

Then, for $U := X \times Y^*, V := X \times X^*, W := Y \times Y^*$, one has

$$E = k^{-1}(H), \quad F = j(E).$$

Proposition 4. *With the preceding assumptions, one has $x^* \in \partial^2 f(\bar{x}, \bar{x}^*)(x^{**})$ if, and only if, for $y^{**} = g'(\bar{x})^{TT}(x^{**})$ and some $y^* \in \partial^2 h(g(\bar{x}), \bar{y}^*)(y^{**})$, one has*

$$x^* = g'(\bar{x})^T(y^*) + \langle x^{**}, \bar{y}^* \circ (D^2 g(\bar{x})(\cdot)) \rangle. \tag{3}$$

Note that for all $u \in X$ one has $D^2 g(\bar{x})(u) \in L(X, Y)$ and $\bar{y}^* \circ (D^2 g(\bar{x})(u)) \in X^*$. If x^{**} is the image of some $x \in X$ through the canonical injection of X into X^{**} , then for $y := g'(\bar{x})(x), y^* \in \partial^2 h(g(\bar{x}), \bar{y}^*)(y)$, one gets

$$x^* = y^* \circ g'(\bar{x}) + \bar{y}^* \circ (D^2 g(\bar{x})(x)),$$

a formula akin the classical formula of the twice differentiable case.

Proof. The derivatives of j and k at $\bar{u} := (\bar{x}, \bar{y}^*)$ are given by

$$\begin{aligned} j'(\bar{x}, \bar{y}^*)(x, y^*) &= (x, \bar{y}^* \circ (D^2 g(\bar{x})(x)) + y^* \circ g'(\bar{x})), \\ k'(\bar{x}, \bar{y}^*)(x, y^*) &= (g'(\bar{x})x, y^*), \end{aligned}$$

so that $k'(\bar{x}, \bar{y}^*)(X \times Y^*) = Y \times Y^*$ and we can apply the last assertion of Proposition 3. Thus, the relation $j'(\bar{u})^T(v^*) = k'(\bar{u})^T(w^*)$ with $v^* := (x^*, -x^{**})$, $w^* := (y^*, -y^{**})$ can be transcribed as

$$\langle (x^*, -x^{**}), (x, \bar{y}^* \circ (D^2g(\bar{x})(x)) + y^* \circ g'(\bar{x})) \rangle = \langle (y^*, -y^{**}), (g'(\bar{x})x, y^*) \rangle.$$

for all $(x, y^*) \in X \times Y^*$. Taking successively $x = 0$ and then $y^* = 0$, we get

$$\begin{aligned} \forall y^* \in Y^* \quad & \langle x^{**}, g'(\bar{x})^T y^* \rangle = \langle y^{**}, y^* \rangle, \\ \forall x \in X \quad & \langle x^*, x \rangle - \langle x^{**}, \bar{y}^* \circ (D^2g(\bar{x})(x)) \rangle = \langle y^*, g'(\bar{x})x \rangle \end{aligned}$$

or $g'(\bar{x})^{TT}(x^{**}) = y^{**}$, $x^* = g'(\bar{x})^T(y^*) + \langle x^{**}, \bar{y}^* \circ (D^2g(\bar{x})(\cdot)) \rangle$. □

5 Coderivative of the Gradient Map of a Moreau Envelope

Let us turn to the coderivative of the gradient map of the Moreau envelope of a paraconvex function f on a Hilbert space X such that for some $a \in \mathbb{R}$, $b \in \mathbb{R}$ one has $f \geq b - a \|\cdot\|^2$. We recalled that there exists some $\bar{r} > 0$ such that for all $r \in]0, \bar{r}[$ the Moreau envelope $e_r f = {}^r f$ is differentiable with gradient $\nabla(e_r f)(x) = (1/r)(x - p_r(x))$ where $p_r(x) \in X$ is such that $e_r(x) = f(p_r(x)) + (1/2r)\|x - p_r(x)\|^2$ and $x - p_r(x) \in r\partial^{\nabla} f(p_r(x))$, where the subgradient $\partial^{\nabla} g(x)$ of g at x is the inverse image of $\partial g(x)$ under the Riesz isomorphism $J : X \rightarrow X^*$. Here we abridge $(e_r f)(x)$ into $e_r(x)$, whereas we write $p_r(x)$ instead of p_r in order to take into account the dependence on x . We have to apply calculus rules for coderivatives. We do not need the full pictures of [24, 26]. A simple lemma is enough. Here we take the coderivatives associated with the directional, the firm (Fréchet), the limiting, and the Clarke subdifferentials, and we simply write ∂ instead of $\partial_D, \partial_F, \partial_L, \partial_C$, respectively.

Lemma 3.

- (a) If $F^{-1} : Y \rightrightarrows X$ is the inverse multimap of $F : X \rightrightarrows Y$, then for all $(y, x) \in F^{-1}$ and all $x^* \in X^*$ one has $y^* \in D^*F^{-1}(y, x)(x^*)$ if and only if $-x^* \in D^*F(x, y)(-y^*)$.
- (b) $u^* \in D^*(-F)(x, -y)(v^*)$ if, and only if, $u^* \in D^*F(x, y)(-v^*)$.
- (c) If $F := I + G$, where $G : X \rightrightarrows X$ and I is the identity map, then for all $v^* \in X^*$ one has $D^*F(x, y)(v^*) = v^* + D^*G(x, y - x)(v^*)$.

Proof. We give the proof in the case the subdifferential is the directional subdifferential ∂_D . The cases of the firm, the limiting, and the Clarke subdifferentials are also simple applications of the definitions.

- (a) We observe that $N(F^{-1}, (y, x)) = (N(F, (x, y)))^{-1}$. Thus, for $x^* \in X^*$ one has $y^* \in D^*F^{-1}(y, x)(x^*)$ if, and only if, $(-x^*, y^*) \in N(F, (x, y))$, if, and only if, $-x^* \in D^*F(x, y)(-y^*)$.
- (b) The equivalence is a consequence of the fact that

$$(u^*, -v^*) \in N(\text{gph}(-F), (x, -y)) \iff (u^*, v^*) \in N(\text{gph}(F), (x, y)).$$

- (c) It is easy to see that $v \in DF(x, y)(u)$ for $(u, v) \in X \times X$ if, and only if, there exists $w \in DG(x, y - x)(u)$ such that $v = u + w$, if, and only if, $v - u \in DG(x, y - x)(u)$. Then $(u^*, -v^*) \in N(F, (x, y))$ if, and only if, for all $u \in X$, $w \in DG(x, y - x)(u)$, one has

$$\langle u^*, u \rangle + \langle -v^*, u + w \rangle \leq 0,$$

if, and only if, $(u^* - v^*, -v^*) \in N(G, (x, y - x))$ or $u^* = v^* + (u^* - v^*) \in v^* + D^*G(x, y - x)(v^*)$. □

Applying the preceding lemma to the map $p_r = (I + r\partial^\nabla f)^{-1}$, we get, since $\nabla e_r(x) = r^{-1}(x - p_r(x))$

$$\begin{aligned} u^* \in D^*\nabla e_r(x)(v^*) &\iff u^* \in r^{-1}(v^* + D^*p_r(x)(-v^*)), \\ w^* \in D^*p_r(x)(-v^*) &\iff v^* \in D^*(I + r\partial^\nabla f)(p_r(x), x)(-w^*) \\ &\iff v^* \in -w^* + rD^*\partial^\nabla f(p_r(x), x)(-w^*). \end{aligned}$$

Therefore, $u^* \in D^*\nabla e_r(x)(v^*)$ if, and only if, for $w^* := ru^* - v^* \in D^*p_r(x)(-v^*)$ one has $v^* + w^* \in rD^*\partial^\nabla f(p_r(x), x)(-w^*)$ or $ru^* \in rD^*\partial^\nabla f(p_r(x), x)(-w^*)$ or $u^* \in D^*\partial^\nabla f(p_r(x), x)(v^* - ru^*)$. We have obtained an expression of $D^*\nabla e_r f$ in terms of $D^*\partial^\nabla f$. We state it as follows.

Proposition 5. *If f is paraconvex around x and such that for some $a \in \mathbb{R}$, $b \in \mathbb{R}$ one has $f \geq b - a \|\cdot\|^2$, then for $r > 0$ small enough, one has $u^* \in D^*\nabla(e_r f)(x)(v^*)$ if, and only if, $u^* \in D^*\partial^\nabla f(p_r(x), x)(v^* - ru^*)$*

$$D^*\nabla(e_r f)(x) = (rI + (D^*\partial^\nabla f(p_r(x), x))^{-1})^{-1}.$$

6 Conclusion

The choice of notation and terminology is not a simple matter. It has some importance in terms of use. The famous example of the quarrel between Leibniz and Newton shows that the choice of the mathematician that is often considered as the first discoverer of a concept is not always the choice that remains in use. For what concerns terminology, in mathematics as in other sciences, both names of scientists and descriptive names are in use. The difference is that mathematicians prefer ordinary names such as “field,” “group,” and “ring” to sophisticated names issued from Greek or Latin. For what concerns notation, there is a trade-off between simplicity and clarity. Both are desirable, but avoiding ambiguity is crucial. Nonetheless, local abuses of notation are often tolerated. Here, because our aim was

limited to a few glances at second-order generalized derivatives, we have avoided a heavy notation involving upper epi-limits and we have often omitted the mention that weak convergence is involved. That is not always suitable.

Let us observe that the use of a generic subdifferential avoids delicate choices of notation for subdifferentials. In his first contributions the author endeavored to use symbols like $f^0(x, \cdot)$, $f^\uparrow(x, \cdot)$ rather than letters like $f^C(x, \cdot)$ to denote generalized derivatives. But the disorder among such symbols led him to adopt a more transparent choice in [38] and elsewhere. Also, he tried to combine the advantages of descriptive terms and authors' names. More importantly, he chose to adopt the simplest notation for simple, fundamental objects such as Fréchet normal cones N_F and Fréchet subdifferentials ∂_F that are used in most proofs rather than affect them with decorations such as \hat{N} or \bar{N} usually kept for completions or limiting constructions. The fact that limiting constructions can be performed by using different convergences comforts that choice.

It would be interesting to decide whether the influences of dominant philosophies (behaviorism in America, Cartesianism, Kantianism, positivism, existentialism, structuralism in Europe) play some role in the different choices present in the mathematical literature. But that question is outside the scope of the present contribution.

References

- [1] Attouch, H.: Variational Convergence for Functions and Operators. Pitman, Boston (1984)
- [2] Bauschke, H.H., Combettes, P.-L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics. Springer, New York (2011)
- [3] Bernard, F., Thibault, L., Zlateva, N.: Prox-regular sets and epigraphs in uniformly convex Banach spaces: various regularities and other properties. *Trans. Amer. Math. Soc.* **363**(4), 2211–2247 (2011)
- [4] Cominetti, R., Correa, R.: A generalized second-order derivative in nonsmooth optimization. *SIAM J. Control Optim.* **28**(4), 789–809 (1990)
- [5] Correa, R., Jofré, A., Thibault, L.: Characterization of lower semicontinuous convex functions. *Proc. Amer. Math. Soc.* **116**(1), 67–72 (1992)
- [6] Correa, R., Jofre, A., Thibault, L.: Subdifferential characterization of convexity. In: Du, D.Z., Qi, L., Womersley, R.S. (eds.) *Recent Advances in Nonsmooth Optimization*, pp. 18–23. World Scientific, Singapore (1995)
- [7] Demyanov, V.F., Rubinov, A.M.: *Constructive Nonsmooth Analysis. Approximation and Optimization*, vol. 7. Peter Lang, Frankfurt am Main, Berlin (1995)
- [8] Demyanov, V.F., Stavroulakis, G.E., Polyakova, L.N., Panagiotopoulos, P.D.: *Quasidifferentiability and Nonsmooth Modelling in Mechanics, Engineering and Economics. Nonconvex Optimization and Its Applications*, vol. 10. Kluwer Academic Publishers, Dordrecht (1996)
- [9] Do, C.N.: Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Trans. Amer. Math. Soc.* **334**(1), 281–301 (1992)
- [10] Hiriart-Urruty, J.-B.: Limiting behaviour of the approximate first order and second order directional derivatives for a convex function. *Nonlinear Anal.* **6**(12), 1309–1326 (1982)
- [11] Hiriart-Urruty, J.-B., Seeger, A.: Calculus rules on a new set-valued second order derivative for convex functions. *Nonlinear Anal.* **13**(6), 721–738 (1989)

- [12] Hiriart-Urruty, J.-B., Seeger, A.: The second-order subdifferential and the Dupin indicatrices of a nondifferentiable convex function. *Proc. London Math. Soc.* (3) **58**(2), 351–365 (1989)
- [13] Huang, L.R., Ng, K.F.: Second-order necessary and sufficient conditions in nonsmooth optimization. *Math. Program.* **66**, 379–402 (1994)
- [14] Huang, L.R., Ng, K.F.: On second-order directional derivatives in nonsmooth optimization. In: Du, D.Z., Qi, L., Womersley, R.S. (eds.) *Recent Advances in Nonsmooth Optimization*, pp. 159–171. World Scientific, Singapore (1995)
- [15] Huang, L.R., Ng, K.F.: On some relations between Chaney’s generalized second order directional derivative and that of Ben-Tal and Zowe. *SIAM J. Control Optim.* **34**(4), 1220–1235 (1996)
- [16] Ioffe, A.: On some recent developments in the theory of second order optimality conditions. In: *Optimization* (Varetz, 1988). *Lecture Notes in Mathematics*, vol. 1405, pp. 55–68. Springer, Berlin (1989)
- [17] Ioffe, A.D.: Variational analysis of a composite function: a formula for the lower second-order epi-derivative. *J. Math. Anal. Appl.* **160**(2), 379–405 (1991)
- [18] Ioffe, A.D.: Theory of subdifferentials. *Adv. Nonlinear Anal.* **1**, 47–120 (2012)
- [19] Ioffe, A.D., Penot, J.-P.: Limiting subhessians, limiting subjets and their calculus. *Trans. Amer. Math. Soc.* **349**, 789–807 (1997)
- [20] Jeyakumar, V., Luc, D.T.: *Nonsmooth Vector Functions and Continuous Optimization*. Springer Optimization and Its Applications, vol. 10. Springer, New York (2008)
- [21] Jourani, A., Thibault, L., Zagrodny, D.: $C^{1,\omega(\cdot)}$ -regularity and Lipschitz-like properties of subdifferential. *Proc. Lond. Math. Soc.* (3) **105**(1), 189–223 (2012)
- [22] Levy, A.B.: Second-order epi-derivatives of composite functionals, in optimization with data perturbations II. *Ann. Oper. Res.* **101**, 267–281 (2001)
- [23] Levy, A.B., Poliquin, R., Thibault, L.: Partial extension of Attouch’s theorem with applications to proto-derivatives of subgradient mappings. *Trans. Amer. Math. Soc.* **347**, 1269–1294 (1995)
- [24] Li, S.J., Penot, J.-P., Xue, X.: Codifferential calculus. *Set-Valued Var. Anal.* **19**(4), 505–536 (2011)
- [25] Michel, Ph., Penot, J.-P.: Second-order moderate derivatives. *Nonlinear Anal. Theor. Methods Appl.* **22**(7), 809–821 (1994)
- [26] Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation I*. Grundlehren der Mathematischen Wissenschaften, vol. 330. Springer, Berlin (2006)
- [27] Mordukhovich, B., Rockafellar, R.T.: Second-order subdifferential calculus with applications to tilt stability in optimization. *SIAM J. Optim.* **22**(3), 953–986 (2012)
- [28] Ndoutoumé, J., Théra, M.: Generalized second-order derivatives of convex functions in reflexive Banach spaces. *Bull. Austr. Math. Soc.* **51**(1), 55–72 (1995)
- [29] Páles, Z., Zeidan, V.: Co-Jacobian for Lipschitzian maps. *Set-Valued Var. Anal.* **18**(1), 57–78 (2010)
- [30] Penot, J.-P.: Second-order derivatives: comparisons of two types of epi-derivatives. In: Oettli, W., Pallaschke, D. (eds.) *Advances in Optimization, Proceedings 6th French-German Conference on Optimization*, Lambrecht, FRG, 1991. *Lecture Notes in Economics and Mathematical Systems*, vol. 382, pp. 52–76. Springer, Berlin (1992)
- [31] Penot, J.-P.: Optimality conditions in mathematical programming and composite optimization. *Math. Program.* **67**(2), 225–246 (1994)
- [32] Penot, J.-P.: Sub-hessians, super-hessians and conjugation. *Nonlinear Anal. Theor. Methods Appl.* **23**(6), 689–702 (1994)
- [33] Penot, J.-P.: Generalized convexity in the light of nonsmooth analysis. In: Duriez, R., Michelot, C. (eds.) *7th French -German Conference on Optimization*, Dijon, July 1994. *Lecture Notes on Economics and Mathematical Systems*, vol. 429, pp. 269–290. Springer, Berlin (1995)
- [34] Penot, J.-P.: Sequential derivatives and composite optimization. *Revue Roumaine Math. Pures Appl.* **40**(5–6), 501–519 (1995)

- [35] Penot, J.-P.: Are generalized derivatives useful for generalized convex functions? In: Crouzeix, J.-P., Martinez-Legaz, J.-E., Volle, M. (eds.) *Generalized Convexity and Generalized Monotonicity*, Marseille, June 1996, pp. 3–59. Kluwer, Dordrecht (1997)
- [36] Penot, J.-P.: On the mean value theorem. *Optimization* **19**(2), 147–156 (1988)
- [37] Penot, J.-P.: Glimpses upon quasiconvex analysis. *ESAIM Proc.* **20**, 170–194 (2007)
- [38] Penot, J.-P.: *Calculus Without Derivatives*. Graduate Texts in Mathematics, vol. 266. Springer, New York (2013)
- [39] Penot, J.-P.: A geometric approach to second-order generalized derivatives (preprint)
- [40] Penot, J.-P.: Directionally limiting subdifferentials and second-order optimality conditions. Appear in *Optimization Letters*. DOI 10.1007/s11590-013-0663-0
- [41] Penot, J.-P.: Towards a new era in subdifferential analysis? in D.H. Bayley et al. (eds), *Computational and Analytical Mathematics*, Springer Proceedings in Mathematics & Statistics 50, (2013) DOI 10.1007/978-1-4614-7621-4_29
- [42] Penot, J.-P., Quang, P.H.: On generalized convexity of functions and generalized monotonicity of set-valued maps. *J. Optim. Theory Appl.* **92**(2), 343–356 (1997)
- [43] Rockafellar, R.T.: First- and second-order epi-differentiability in nonlinear programming. *Trans. Amer. Math. Soc.* **307**, 75–107 (1988)
- [44] Rockafellar, R.T.: Second-order optimality conditions in nonlinear programming obtained by the way of pseudo-derivatives. *Math. Oper. Res.* **14**, 462–484 (1989)
- [45] Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Grundlehren der Mathematischen Wissenschaften, vol. 317. Springer, Berlin (1998)
- [46] Warga, J.: Fat homeomorphisms and unbounded derivate containers. *J. Math. Anal. Appl.* **81**(2), 545–560 (1981)
- [47] Yang, X.Q.: On relations and applications of generalized second-order directional derivatives. *Nonlin. Anal. Theor. Methods Appl.* **36**(5), 595–614 (1999)