INFIMAL CONVOLUTIONS AND LIPSCHITZIAN PROPERTIES OF SUBDIFFERENTIALS FOR PROX-REGULAR FUNCTIONS IN HILBERT SPACES

MIROSLAV BAČÁK, JONATHAN M. BORWEIN, ANDREW EBERHARD, AND BORIS S. MORDUKHOVICH

ABSTRACT. In this paper we study infimal convolutions of extended-real-valued functions in Hilbert spaces paying a special attention to the rather broad and remarkable class of prox-regular functions. Such functions have been well recognized as highly important in many aspects of variational analysis and its applications in both finite-dimensional and infinite-dimensional settings. Based on advanced variational techniques, we discover some new subdifferential properties of infimal convolutions and apply them to the study of Lipschitzian behavior of subdifferentials for prox-regular functions in Hilbert spaces. It is shown, in particular, that the fulfillment of a natural Lipschitz-like property for (set-valued) subdifferentials of prox-regular functions forces such functions, under weak assumptions, actually to be locally smooth with single-valued subdifferentials reduced to Lipschitz continuous gradient mappings.

1. INTRODUCTION

This paper is mainly devoted to the study of *infimal convolutions* of extended-real-valued functions in Hilbert spaces, with our particular attention given to the so-called *prox-regular* functions introduced in 1996 by Poliquin and Rockafellar [27] in the context of finite-dimensional spaces. Since that time, this remarkable class of functions has been demonstrated to be most useful in many aspects of variational theory and its applications. On one hand, the construction of prox-regularity is able to guarantee many desirable properties of such functions and, on the other hand, it is broad enough to accommodate various important classes of functions including lower semicontinuous convex functions, strongly amenable functions, or lower-C^2 (i.e., continuous locally para-convex) functions [27]. More recently Bernard and Thibault [6, 7, 8] generalized the concept of prox-regularity to infinite-dimensional spaces and proved several fundamental results in infinite-dimensional settings. The aim of our paper is to continue with further development of prox-regularity in Hilbert spaces along those lines, namely to show new subdifferentiability properties of prox-regular functions. Moreover, while extensions exist to uniformly convex Banach space [9], the results below seem most useful and quite possibly valid only in Hilbert spaces.

2000 *Mathematics Subject Classification.* 49J52, 46C05.

Key words and phrases. Subdifferentials, Lipschitz continuity, infimal convolutions, prox-regular functions, prox-bounded functions, set-valued mappings.

The paper is dedicated to Hedi Attouch on the occasion of his sixtieth birthday. This research was supported in part by the ARC Discovery grants DP0664423 and DP0987445. The research of the forth author was also partly supported by the US National Science Foundation under grant DMS-0603846.

The crucial tools of our analysis involve the aforementioned infimal convolutions known also as *regularizations* or as *Moreau/proximal envelopes*. We study infimal convolutions of lower semicontinuous functions defined on Hilbert spaces and obtain new properties and characterizations using appropriate subdifferential constructions. Then we establish more specific and stronger results for infimal convolutions of prox-regular functions. It is worth mentioning that our techniques are significantly different from those developed earlier (see, e.g., [6, 7, 8, 19, 17, 14, 27] and the references therein). In particular, we completely avoid employing the so-called f-attentive localization of subdifferentials for the functions under consideration.

Based on the infimal convolution techniques and results developed herein, we prove in the general Hilbert space setting that the underlying subdifferential mapping for a prox-regular function turns out to locally *single-valued* and *Lipschitz continuous* under rather mild and seemingly natural requirements. This fact has a number of interesting consequences in variational theory and applications some of which are discussed in the paper.

The rest of the paper is organized as follows. Section 2 contains basic definitions and preliminaries widely used in the subsequent material. For the central result of Theorem 2.3 (a striking characterization of $C^{1,1}$ functions via para-convexity and para-concavity) we give a new proof based in second-order differentiability.

Section 3 concerns subdifferential properties of infimal convolutions for lower semicontinuous functions in Hilbert spaces. Some of the results presented here are known in finite dimensions while others are new in both finite-dimensional and infinite-dimensional settings.

Sections 4 and 5 contain the main results of the paper. In particular, Theorem 4.9 establishes the uniform prox-regularity of infimal convolutions of prox-regular functions in Hilbert spaces with computing the corresponding moduli. Theorem 4.11 justifies a local $C^{1,1}$ property for infimal convolutions of such functions. Finally, Theorem 5.4 proves the aforementioned local single-valuedness and Lipschitz continuity of the subdifferential mappings for prox-regular functions on Hilbert spaces.

Our notation is basically conventional in the area of variational analysis; see, e.g., [11, 25, 30] and Section 2 for more details. Recall that, given a set-valued mapping $F : H \rightrightarrows H$ from a Hilbert space to itself, the *Painlevé-Kuratowski outer/upper limit* of $F(x)$ as $x \to \overline{x}$ is defined by

$$\operatorname*{Lim\,sup}_{x\to\overline{x}} F(x) := \Big\{x^* \in H \Big| \, \exists \text{ sequences } x_n \to \overline{x},\ x_n^* \in F(x_n),\ x_n^* \xrightarrow{w} x^*\Big\},$$

where the symbol $\xrightarrow{w}$ signifies the sequential convergence in the *weak* topology of H.

2. Some Preliminaries with Alternative Proofs

Let H be a real *Hilbert space* endowed with the *inner product* $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{R}$. The induced *norm* is denoted by $\|\cdot\|$ and an *open ball* centered at $x \in H$ by $B_\delta(x) := \big\{y \in H \big| \, \|x-y\| < \delta\big\}$. Given a set $A \subset H$, denote the *norm closure* of A by $\overline{A}$ and *weakly closed convex hull* by $\overline{\text{co}}\, A$. By the *domain of a function* $f : H \to (-\infty, \infty]$ we mean the set $\text{dom}\, f := \{x \in H |\ f(x) < \infty\}$, whereas the *domain of a multifunction/set-valued mapping* $F : H \rightrightarrows H$ is the set $\text{dom}\, F := \{x \in H |\ F(x) \neq \emptyset\}$. The *Gâteaux derivative* of a function $f : H \to \mathbb{R}$ at a point

$x \in H$ is denoted by $\nabla f(x)$, and the *derivative* at $x \in H$ *in a direction* $u \in H$ by

$$f'(x;u) := \lim_{t \downarrow 0} \frac{f(x+tu) - f(x)}{t}.$$

The symbol δ_Ω stands for the *indicator function* of a set $\Omega \subset H$, that is, $\delta_\Omega(x) := 0$ for $x \in \Omega$ and ∞ otherwise. If $f : H \to (-\infty, \infty]$ is a function and $(x_n) \subset H$ a sequence, we define *f-attentive convergence* of (x_n) by

$$x_n \to^f x \qquad \text{if} \quad x_n \to x \;\text{ and }\; f(x_n) \to f(x).$$

We say that a function $f : H \to (-\infty, \infty]$ is *lower semicontinuous* around $x \in H$ if there exist $\alpha, \beta > 0$ such that f is lower semicontinuous on

$$\left\{ y \in H \middle| \; \|x - y\| < \alpha, \; f(y) < f(x) + \beta \right\}.$$

Lower semicontinuity is often be abbreviated to *lsc* in what follows.

Definition 2.1. *We say that a function $f : H \to (-\infty, \infty]$ is:*

(i) LOCALLY PARA-CONVEX *around $x \in H$ if the function $f + \frac{\lambda}{2}\|\cdot\|^2$ is convex and continuous on $B_\delta(x)$, for some $\delta > 0$, and $\lambda \geq 0$;*

(ii) LOCALLY $C^{1,1}$ *(known also as $C^{1,+}$) around $x \in H$ if the derivative ∇f exists everywhere inside of $B_\delta(x)$, for some $\delta > 0$, and the derivative mapping $y \mapsto \nabla f(y)$ is Lipschitz continuous on $B_\delta(x)$;*

(iii) LOCALLY DIRECTIONALLY $C^{1,1}$ *around $x \in H$ if ∇f exists on $B_\delta(x)$, for some $\delta > 0$, and there is $\lambda > 0$ such that for all $u, v \in B_\delta(x)$ we have*

$$|\langle \nabla f(u) - \nabla f(v), u - v \rangle| \leq \lambda \|u - v\|^2.$$

A function $f : H \to \mathbb{R} \cup \{-\infty\}$ is LOCALLY PARA-CONCAVE *around $x \in H$ if the function $-f$ is locally para-convex around $x \in H$.*

The following remarkable result was proved in [20, Corollary 2].

Proposition 2.2. *Let $f : H \to \mathbb{R}$ be a Gâteaux differentiable function, and let $\lambda > 0$. Then f is locally directionally $C^{1,1}$ with Lipschitz constant λ if and only if f is locally $C^{1,1}$ with Lipschitz constant λ.*

To the best of our knowledge, Hiriart-Urruty and Plazanet [20] have been the first to observe, along with Proposition 2.2, that a real-valued function is locally $C^{1,1}$ if it is simultaneously locally para-convex and para-concave. A related observation was implicitly used in [22]. We now provide, employing some ideas from [15], an alternative proof indicating the new lines of connection of this set of results to *second-order differentiability* via the classical Alexandrov theorem; see [10] and the references therein.

Theorem 2.3. *A function $f : H \to \mathbb{R}$ is locally $C^{1,1}$ around $\overline{x} \in H$ if and only if it is simultaneously locally para-convex and locally para-concave around $\overline{x}$.*

Proof. Suppose f is $C^{1,1}$ in $B_\delta(\overline{x})$. Then ∇f is Lipschitz in $B_\delta(\overline{x})$ with a Lipschitz constant $\lambda > 0$. This gives, for $u, y \in B_\delta(\overline{x})$ and $\alpha \in [0,1]$, that

$$\begin{aligned} &|\langle \nabla f(\alpha y + (1-\alpha)u) - \nabla f(u), y - u \rangle| \\ &\qquad = \frac{1}{\alpha} |\langle \nabla f\left(\alpha(y-u) + u\right) - \nabla f(u), \alpha(y-u) \rangle| \\ &\qquad \leq \alpha \lambda \|y - u\|^2, \end{aligned}$$

and so we have the relationships

$$\begin{aligned} f(y) - f(u) - \langle \nabla f(u), y-u\rangle & \\ &= \int_0^1 \langle \nabla f\left(\alpha y + (1-\alpha)u\right) - \nabla f(u), y-u\rangle \,\mathrm{d}\alpha \\ &\geq -\lambda \|y-u\|^2 \int_0^1 \alpha \,\mathrm{d}\alpha = -\frac{\lambda}{2}\|y-u\|^2 \\ &= -\frac{\lambda}{2}\left(\|y\|^2 - \|u\|^2 + 2\|u\|^2 - 2\langle u, y\rangle\right). \end{aligned}$$

The latter implies the gradient/subgradient inequality of convex analysis

$$f(y) + \frac{\lambda}{2}\|y\|^2 - \left(f(u) + \frac{\lambda}{2}\|u\|^2\right) \geq \langle \nabla f(u) + \lambda u, y-u\rangle$$

thus verifies that the function $f + \frac{\lambda}{2}\|\cdot\|^2$ is convex inside of $B_\delta(\overline{x})$. It is clear, by a similar argument, that $f - \frac{\lambda}{2}\|\cdot\|^2$ is concave in $B_\delta(\overline{x})$ when f is $C^{1,1}$ inside the ball $B_\delta(\overline{x})$. This justifies the "only if" part of the theorem.

For the converse, note that we may take $\lambda > 0$ sufficiently large so that both $f + \frac{\lambda}{2}\|\cdot\|^2$ is convex and $f - \frac{\lambda}{2}\|\cdot\|^2$ is concave in some neighborhood $B_\delta(\overline{x})$. Then at each point $x \in B_\delta(\overline{x})$ there exists a subgradient $v(x) \in \partial\left(f + \frac{\lambda}{2}\|\cdot\|^2\right)(x)$ and a supergradient $w(x) \in \partial\left(f - \frac{\lambda}{2}\|\cdot\|^2\right)(x)$. Thus we have

$$\langle v(x), u-x\rangle - \frac{\lambda}{2}\left(\|u\|^2 - \|x\|^2\right) \leq f(u) - f(x) \leq \langle w(x), u-x\rangle + \frac{\lambda}{2}\left(\|u\|^2 - \|x\|^2\right).$$

To proceed with proving $f \in C^{1,1}$ around $\overline{x}$, let us show first that the gradient $\nabla f(x)$ exists. Use $u = x + tz$ in the above inequalities to obtain

$$\langle v(x) - \lambda x, z\rangle - \frac{\lambda}{2}t\|z\|^2 \leq \frac{1}{t}\left(f(x+tz) - f(x)\right) \leq \langle w(x) + \lambda x, z\rangle + \frac{\lambda}{2}t\|z\|^2.$$

By letting $t \downarrow 0$, we get the inequalities

$$\langle v(x) - \lambda x, z\rangle \leq f'(x;z) \leq \langle w(x) + \lambda x, z\rangle \text{ for all } z.$$

The linearity of the functions in the upper and lower bounds in z implies that

$$v(x) - \lambda x = w(x) + \lambda x := \nabla f(x).$$

It remains to prove the Lipschitz continuity of ∇f on $B_\delta(x)$. For any given $x, x' \in B_\delta(\overline{x})$, consider the convex function of one variable

$$\alpha \mapsto f(\alpha x + (1-\alpha)x'),$$

which is differentiable with derivative

$$f'\left(\alpha x + (1-\alpha)x'\right) = \langle \nabla f\left(\alpha x + (1-\alpha)x'\right), x - x'\rangle.$$

From the classical Alexandrov theorem on the real line we know that for each $\lambda > 0$ the real-valued function

$$\alpha \mapsto f(\alpha x + (1-\alpha)x') \pm \frac{\lambda}{2}\|\alpha x + (1-\alpha)x'\|^2$$

is twice differentiable almost everywhere, with respect to Lebesgue measure, on the interval $[0,1]$. At each point of second-order differentiability $\alpha \in S_+ \subset [0,1]$ of the function $\alpha \mapsto f(\alpha x + (1-\alpha)x') + \frac{\lambda}{2}\|\alpha x + (1-\alpha)x'\|^2$ we have

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2}\left(f(\alpha x + (1-\alpha)x') + \frac{\lambda}{2}\|\alpha x + (1-\alpha)x'\|^2\right) \geq 0.$$

This gives us the estimate

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2} f(\alpha x + (1-\alpha)x') \geq -\lambda\|x - x'\|^2.$$

Similarly, for every point of second-order differentiability $\alpha \in S_- \subset [0,1]$ of the function $\alpha \mapsto f(\alpha x + (1-\alpha)x') - \frac{\lambda}{2}\|\alpha x + (1-\alpha)x'\|^2$ we have

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2} f(\alpha x + (1-\alpha)x') \leq \lambda\|x - x'\|^2.$$

Then for $\alpha \in S_+ \cap S_-$ the following inequality holds:

$$\left\|\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2} f(\alpha x + (1-\alpha)x')\right\| \leq \lambda\|x - x'\|^2.$$

Since $S_+ \cap S_-$ is a set of full Lebesgue measure on $[0,1]$, we get

$$\begin{aligned}\|\langle \nabla f(x) - \nabla f(x'), x - x'\rangle\| &\leq \int_0^1 \left\|\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2} f(\alpha x + (1-\alpha)x')\right\| \mathrm{d}\alpha \\ &\leq \lambda\|x - x'\|^2 \int_0^1 \mathrm{d}\alpha = \lambda\|x - x'\|^2\end{aligned}$$

and thus complete the proof of the theorem. □

We will see in Theorem 4.11 that the result of Theorem 2.3 allows us to provide a direct proof of the fact that every lsc prox-regular function defined on a Hilbert space admits a $C^{1,1}$ infimal convolution. Let us now recall the definition of the latter construction, which plays a crucial role in this paper.

Infimal convolutions. Given $\lambda > 0$ and $f : H \to (-\infty, \infty]$, define the *infimal convolution* of f at $x \in H$ by

$$f_\lambda(x) := \inf_{u \in H}\left(f(u) + \frac{1}{2\lambda}\|x - u\|^2\right) \tag{1}$$

and the corresponding *proximal mapping* by

$$P_\lambda(x) := \operatorname*{arg\,max}_{u \in H}\left(f(u) + \frac{1}{2\lambda}\|x - u\|^2\right). \tag{2}$$

The next simple proposition is useful in what follows.

Proposition 2.4. *For any function $f : H \to (-\infty, \infty]$ the infimal convolution* (1) *is para-concave around each point $x \in \operatorname{dom} f$.*

Proof. Applying the infimal convolution (1) to any (nonconvex) function f, we get by definition the following equalities:

$$\begin{aligned}-\left(f_\lambda(x) - \frac{1}{2\lambda}\|x\|^2\right) &= -\left[\inf_{u \in H}\left(f(u) + \frac{1}{2\lambda}\|x - u\|^2\right) - \frac{1}{2\lambda}\|x\|^2\right] \\ &= \sup_{u \in H}\left[\frac{1}{\lambda}\langle x, u\rangle - \left(\frac{1}{2\lambda}\|u\|^2 + f(u)\right)\right].\end{aligned}$$

The latter function is a supremum of affine functions, and so it is convex. Thus the function $f_\lambda - \frac{1}{2\lambda}\|\cdot\|^2$ is concave, and consequently f_λ is para-concave. □

It is worth mentioning here that when f is prox-regular (see Section 4), then the function $f_\lambda + \frac{1}{2\lambda}\|\cdot\|^2$ is convex for λ sufficiently small. We can also show that f_λ is a locally Lipschitzian function for λ sufficiently small, and furthermore the Lipschitz constant can be taken of order $O(\lambda^{-1})$.

3. Subdifferential Properties of Infimal Convolutions

In this section we establish some new properties of infimal convolutions such as are needed for deriving the main results in Sections 4 and 5. Our close attention is paid here to so-called *prox-bounded* functions $f : H \to (-\infty, \infty]$ (see [30]), which can be equivalently described as follows: there is $r \in \mathbb{R}$ such that the function $f + \frac{r}{2}\|\cdot\|^2$ is bounded from below. The latter is the same as the assumption that f can be quadratically minorized, by a quadratic function of the form $\alpha - \frac{r}{2}\|\cdot\|^2$. Thus a sufficient condition for $f_\lambda > -\infty$ is that $\lambda < (\max\{0, r\})^{-1}$, and then $P_\lambda(x) \neq \emptyset$ for the proximal mapping (2). The infimum of all such r is denoted by $r(f)$. It is possible that $r(f) < 0$, and so we define the nonnegative modulus $\overline{r}(f) := \max\{r(f), 0\}$. The number $\lambda_f := (\overline{r}(f))^{-1}$ is called the *proximal threshold* for f, with the convention that $1/0 := \infty$. Therefore, for $r(f) < 0$ we get that $f_\lambda > -\infty$ whenever $\lambda > 0$. It is well known that the family $\{f_\lambda\}_{\lambda>0}$ converges to f monotonically (upward) pointwise, and hence it epi-converges to f [10, 30]. It also converges uniformly on bounded sets when f is continuous and real-valued.

Let us next define the notions of generalized differentiability widely used in this and subsequent sections; see [11, 12, 14, 25, 30, 31] for more details and references.

Definition 3.1. *Consider $f : H \to (-\infty, \infty]$ and $x \in \operatorname{dom} f$.*

(i) *We call $p \in H$ a* PROXIMAL SUBGRADIENT *of f at x if there is $r \geq 0$ with*

$$f(x') \geq f(x) + \langle p, x' - x\rangle - \frac{r}{2}\|x' - x\|^2$$

for any x' from some neighborhood of x. The PROXIMAL SUBDIFFERENTIAL *$\partial_{\mathrm{p}} f(x)$ of f at x is the collection of all proximal subgradients of f at x.*

(ii) *The (basic, limiting, Mordukhovich)* SUBDIFFERENTIAL *of f at x is*

$$\partial f(x) = \operatorname*{Lim\,sup}_{x' \to^f x} \partial_{\mathrm{p}} f(x') := \left\{ \text{weak} - \lim v_n \,\middle|\, v_n \in \partial_{\mathrm{p}} f(x_n),\ x_n \to^f x \right\}.$$

(iii) *Let $p \in H$ and Q be a symmetric bilinear form on H. A pair (p, Q) belongs to the* SUBJET *of f at x if there exists $\delta > 0$ such that for all $x' \in B_\delta(x)$ we have the inequality*

$$f(x') \geq f(x) + \langle p, x' - x\rangle + \frac{1}{2} Q\left(x' - x, x' - x\right) + o\left(\|x' - x\|^2\right).$$

In this case we write $(p, Q) \in \partial^{2,-} f(x)$.

It follows from the definitions that $p \in \partial_{\mathrm{p}} f(x)$ if and only if $(p, Q) \in \partial^{2,-} f(x)$ for some symmetric bilinear form Q on H.

Definition 3.2. *Let $f : H \to \mathbb{R}$ be locally Lipschitzian around $x \in H$. The (Clarke)* GENERALIZED DIRECTIONAL DERIVATIVE *of f at x in the direction $u \in H$, denoted* *by $f^\circ(x;u)$, is defined as follows:*

$$f^\circ(x;u) := \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+tu)-f(y)}{t},$$

where $y \in H$ and $t > 0$. The (Clarke) GENERALIZED GRADIENT *of f at x is*

$$\partial_{\rm C} f(x) := \left\{ z \in H \,\middle|\, \langle z,u\rangle \le f^\circ(x;u) \;\; \text{for all} \;\; u \in H \right\}.$$

Observe further that $f^\circ(x;u) = \sigma\left(\partial_{\rm C} f(x), u\right) := \sup\left\{\langle z,u\rangle \,|\, z \in \partial_{\rm C} f(x)\right\}$. When f is locally Lipschitzian around $x \in H$, we have $\partial_{\rm C} f(x) = \overline{\rm co}\,\partial f(x)$; see, e.g., [25, Theorem 3.57]. Thus we have for locally Lipschitzian functions that the $\partial f(x) = \partial_{\rm p} f(x)$ implies that $\partial f(x) = \partial_{\rm C} f(x)$ by the convexity of $\partial_{\rm p} f(x)$ and the weak closedness of $\partial f(x)$.

The next lemma is a slight modification of [13, Corollary 2.1].

Lemma 3.3. *Let $f : H \to (-\infty,\infty]$ be a lsc function with $\operatorname{dom} f \neq \emptyset$, let $\partial^\bullet f$ stand for either the basic subdifferential ∂f or the generalized gradient $\partial_{\rm C} f$ (in which case we assume f to be locally Lipschitz), and let U be a convex subset of H. Then the following characterizations hold:*

(i) *f is convex on U if and only if $\partial^\bullet f$ is monotone in U, i.e.,*

$$\langle z_1^* - z_2^*, x_1 - x_2\rangle \ge 0$$

for all $z_i^* \in \partial^\bullet f(x_i)$ and $x_i \in U \cap \operatorname{dom} \partial^\bullet f$, for $i = 1,2$.

(ii) *f is para-convex on U if and only if $\partial^\bullet f$ is hypomonotone, i.e., for each $x \in \operatorname{dom} \partial^\bullet f \cap U$ there are $\varepsilon > 0$ and $r \ge 0$ such that $B_\varepsilon(x) \subset U$ and*

$$\langle z_1^* - z_2^*, x_1 - x_2\rangle \ge -r\|x_1 - x_2\|^2$$

for all $z_i^* \in \partial^\bullet f(x_i)$ and $x_i \in B_\varepsilon(x) \cap \operatorname{dom} \partial^\bullet f$, for $i = 1,2$.

The next result provides an extension to the Hilbert space setting of the fact observed in [14] in finite-dimensional spaces.

Lemma 3.4. *Assuming that $f : H \to (-\infty,\infty]$ is lsc and prox-bounded, we have that for all $x \in \operatorname{dom} f$ and $\lambda > 0$ sufficiently small the inclusion $(p,Q) \in \partial^{2,-} f_\lambda(x)$ implies the fulfilment of both inclusions $(p,Q) \in \partial^{2,-} f(x - \lambda p)$ and*

$$f(x - \lambda p) = f_\lambda(x) - \frac{\lambda}{2}\|p\|^2.$$

Proof. As in [2, Proposition 1.2a], given $\lambda > 0$ sufficiently small and $x \in \operatorname{dom} f$, there is $\rho > 0$ such that

$$f_\lambda(x) = \inf_{y \in B_\rho(x)} f(y) + \frac{1}{2\lambda}\|x - y\|^2.$$

Choose a sequence $\varepsilon_i \downarrow 0$ and pick $y_i \in B_\rho(x)$ such that

$$f_\lambda(x) + \varepsilon_i \ge f(y_i) + \frac{1}{2\lambda}\|y_i - x\|^2. \tag{3}$$

The inclusion $(p,Q) \in \partial^{2,-} f_\lambda(x)$ means there exists $\delta > 0$ such that

$$\begin{aligned} f_\lambda(x') &\ge f_\lambda(x) + \langle p, x' - x\rangle + \frac{1}{2}Q(x'-x, x'-x) + o(\|x'-x\|^2) \\ &\ge f(y_i) + \frac{1}{2\lambda}\|x - y_i\|^2 - \varepsilon_i + \langle p, x'-x\rangle + \frac{1}{2}Q(x'-x, x'-x) + o(\|x'-x\|^2). \end{aligned} \tag{4}$$

for $x' \in B_\delta(x)$. Clearly, for all $\xi \in H$ we have

$$f(\xi) + \frac{1}{2\lambda}\|x' - \xi\|^2 \geq f_\lambda(x'). \tag{5}$$

Let us next demonstrate that $y_i \to y := x - \lambda p$ in the norm topology. Combine (4) with (5), where we take $\xi = y_i$ and $x' = x + \alpha(x - y_i)/\lambda - \alpha p$ with $\alpha < 0$ such that $\|\alpha(x - y_i)/\lambda - \alpha p\| < \delta$, to get the estimate

$$\frac{1}{2\lambda}\left\|x - y_i + \alpha\frac{x - y_i}{\lambda} - \alpha p\right\|^2 \geq \frac{1}{2\lambda}\|x - y_i\|^2 + \left\langle p, \alpha\frac{x - y_i}{\lambda} - \alpha p\right\rangle + o(\alpha^2) - \varepsilon_i,$$

which upon simplification yields

$$\varepsilon_i \geq \left\|\frac{x - y_i}{\lambda} - p\right\|^2 + o(\alpha).$$

Now passing to the limit as $\alpha \uparrow 0$ shows that

$$\lambda^2 \varepsilon_i \geq \|(x - \lambda p) - y_i\|^2,$$

and thus $y_i \to y$. Since f is lsc, we can pass to the limit in (4) as $y_i \to y$ and get

$$f_\lambda(x') \geq f(y) + \frac{1}{2\lambda}\|x - y\|^2 + \langle p, x' - x\rangle + \frac{1}{2}Q(x' - x, x' - x) + o\left(\|x' - x\|^2\right). \tag{6}$$

Substituting $x' = \xi - y + x$ with $\|\xi - y\| < \delta$ into (5) and (6) yields

$$f(\xi) \geq f(y) + \langle p, \xi - y\rangle + \frac{1}{2}Q(\xi - y, \xi - y) + o(\|\xi - y\|^2),$$

which in turn implies that

$$(p, Q) \in \partial^{2,-} f(y).$$

Finally, from (3) we have $f_\lambda(x) = f(y) + \frac{1}{2\lambda}\|x - y\|^2$ and thus arrive at

$$f_\lambda(x) - \frac{\lambda}{2}\|p\|^2 = f(y) + \frac{1}{2\lambda}\|x - y\|^2 - \frac{\lambda}{2}\|p\|^2 = f(y) + \frac{\lambda}{2}\|p\|^2 - \frac{\lambda}{2}\|p\|^2 = f(y),$$

which completes the proof of the lemma. □

Remark 3.5. A consequence of Lemma 3.4 is the observation that, whenever $p \in \partial_{\mathrm{p}} f_\lambda(x) \neq \emptyset$, the infimum in the infimal convolution is attained. Indeed, we then have the equality

$$f_\lambda(x) = f(x - \lambda p) + \frac{1}{2\lambda}\|x - (x - \lambda p)\|^2. \tag{7}$$

The next lemma allows us, in particular, to prove the reverse implication to (7).

Lemma 3.6. *For any function $f : H \to (-\infty, \infty]$, any $\lambda > 0$ and $x, p \in H$ we have*

$$\left(f - \langle p, \cdot\rangle\right)_\lambda (x) = f_\lambda(x + \lambda p) - \langle p, x\rangle - \frac{\lambda}{2}\|p\|^2. \tag{8}$$

Moreover, $f(x) + \frac{\lambda}{2}\|p\|^2 = f_\lambda(x + \lambda p)$ iff $\left(f - \langle p, \cdot\rangle\right)_\lambda (x) = f(x) - \langle p, x\rangle$ iff

$$f(w) \geq f(x) + \langle p, w - x\rangle - \frac{1}{2\lambda}\|x - w\|^2 \qquad \textit{for all } w \in H, \tag{9}$$

which in turn implies that $p \in \partial_{\mathrm{p}} f(x)$.

Proof. By direct calculation we get the relationships

$$\begin{aligned}(f-\langle p,\cdot\rangle)_\lambda(x) &= \inf_{w\in H}\left(f(w)-\langle p,w\rangle+\frac{1}{2\lambda}\|w-x\|^2\right)\\ &= \inf_{w\in H}\left(f(w)+\frac{1}{2\lambda}\left(\|\lambda p\|^2-2\langle\lambda p,w-x\rangle+\|w-x\|^2\right)\right)\\ &\quad -\langle p,x\rangle-\frac{\lambda}{2}\|p\|^2\\ &= \inf_{w\in H}\left(f(w)+\frac{1}{2\lambda}\|w-(x+\lambda p)\|^2\right)-\langle p,x\rangle-\frac{\lambda}{2}\|p\|^2\\ &= f_\lambda(x+\lambda p)-\langle p,x\rangle-\frac{\lambda}{2}\|p\|^2,\end{aligned}$$

which justify (8). Now suppose that $f(x)+\frac{\lambda}{2}\|p\|^2=f_\lambda(x+\lambda p)$ and deduce from (8) the equalities

$$f_\lambda(x+\lambda p)=(f-\langle p,\cdot\rangle)_\lambda(x)+\langle p,x\rangle+\frac{\lambda}{2}\|p\|^2=f(x)+\frac{\lambda}{2}\|p\|^2,$$

which give $(f-\langle p,\cdot\rangle)_\lambda(x)=f(x)-\langle p,x\rangle$. By definition (1) of the infimal convolution we have, for all $w\in H$, that

$$f(x)-\langle p,x\rangle\le f(w)-\langle p,w\rangle+\frac{1}{2\lambda}\|x-w\|^2,$$

and thus (9) holds. The latter is clearly equivalent to $(f-\langle p,\cdot\rangle)_\lambda(x)=f(x)-\langle p,x\rangle$. By using finally (8), we arrive at

$$f(x)-\langle p,x\rangle=f_\lambda(x+\lambda p)-\langle p,x\rangle-\frac{\lambda}{2}\|p\|^2,$$

which therefore justifies

$$f_\lambda(x+\lambda p)=f(x)+\frac{\lambda}{2}\|p\|^2$$

and completes the proof of the lemma. □

The next result concerning prox-bounded functions can be found in [17] in finite dimensions; herein we extend it to the Hilbert space setting.

Lemma 3.7. *Suppose that $f:H\to(-\infty,\infty]$ is lsc and prox-bounded. Let $\overline{x}\in\operatorname{dom} f$ and $0\in\partial_{\mathrm{p}}f(\overline{x})$. Then there exists a nonnegative number r such that*

$$f(x)\ge f(\overline{x})-\frac{r}{2}\|x-\overline{x}\|^2\qquad\text{for all }\ x\in H. \tag{10}$$

Proof. Since $0\in\partial_{\mathrm{p}}f(\overline{x})$ there exist r_1 and $\delta>0$ ensuring that

$$f(x)\ge f(\overline{x})-\frac{r_1}{2}\|x-\overline{x}\|^2\qquad\text{for all }\ x\in B_\delta(\overline{x}).$$

From the prox-boundedness of f we find $\overline{r}>0$ and $\alpha\in\mathbb{R}$ such that

$$\begin{aligned}f(x)&\ge\alpha-\frac{\overline{r}}{2}\|x\|^2\\ &=\alpha-\frac{\overline{r}}{2}\|x-\overline{x}\|^2+\frac{\overline{r}}{2}\|\overline{x}\|^2+\overline{r}\langle x-\overline{x},\overline{x}\rangle\\ &\ge\alpha-\frac{\overline{r}}{2}\|x-\overline{x}\|^2+\frac{\overline{r}}{2}\|\overline{x}\|^2-\overline{r}\|\overline{x}\|\cdot\|x-\overline{x}\|\end{aligned}$$

for all $x \in H$. Now suppose that $x \notin B_\delta(\overline{x})$ and hence $\|x-\overline{x}\| \geq \delta$. Then

$$f(x) \geq \alpha - \frac{\overline{r}}{2}\|x\|^2 \geq \alpha + \frac{\overline{r}}{2}\|\overline{x}\|^2 - \frac{\overline{r}}{2}\left(\frac{2}{\delta}\|\overline{x}\| + 1\right)\|x-\overline{x}\|^2.$$

Letting $k := \frac{2}{\delta}\|\overline{x}\|$, we find $r_2 \geq 0$ sufficiently large so that

$$\alpha + \frac{\overline{r}}{2}\|\overline{x}\|^2 - \frac{\overline{r}(1+k)}{2}\|x-\overline{x}\|^2 \geq f(\overline{x}) - \frac{r_2}{2}\|x-\overline{x}\|^2$$

for all $x \notin B_\delta(\overline{x})$. Hence

$$f(x) \geq f(\overline{x}) - \frac{r_2}{2}\|x-\overline{x}\|^2$$

for all $x \notin B_\delta(\overline{x})$. Putting $r := \max\{r_1, r_2\}$ finishes the proof. □

Having established in Lemma 3.7 the existence of a value of $r \geq 0$ for which (10) holds, we take the smallest r satisfying (10) and denote this number by $\overline{r}(f,\overline{x})$. It follows from the definitions of $\overline{r}(f,\overline{x})$ and of the nonnegative modulus of prox-boundedness $\overline{r}(f)$ given at the beginning of this section that $\overline{r}(f,\overline{x}) \geq \overline{r}(f)$.

The next result clarifies relationships between the prox-boundedness and proximal subdifferential of f on its domain.

Proposition 3.8. *Suppose that $f : H \to (-\infty,\infty]$ is lsc and prox-bounded. Take any $\overline{x} \in \operatorname{dom} f, v \in \partial_{\mathrm{p}} f(\overline{x})$, and $\mu \in (0,1/r)$, where $r > 0$ is such that*

$$f(x) \geq f(\overline{x}) + \langle v, x-\overline{x}\rangle - \frac{r}{2}\|x-\overline{x}\|^2 \qquad \text{for all } x \in H. \tag{11}$$

Then we have the inclusion $v \in \partial_{\mathrm{p}} f_\mu(\overline{x}+\mu v)$.

Proof. Since f is prox-bounded and $v \in \partial_{\mathrm{p}} f(\overline{x})$, we know that there exists $r > 0$ satisfying (11). For all $x \in H$ and $\mu \in (0,1/r)$ it follows from (11) that

$$\left(f - \langle v,\cdot\rangle\right)(x) + \frac{1}{2\mu}\|x-\overline{x}\|^2 \geq \left(f - \langle v,\cdot\rangle\right)(\overline{x}).$$

Hence we get the inequalities

$$f(\overline{x}) - \langle v,\overline{x}\rangle \geq (f - \langle v,\cdot\rangle)_\mu(\overline{x}) \geq f(\overline{x}) - \langle v,\overline{x}\rangle,$$

and thus $(f - \langle v,\cdot\rangle)_\mu(\overline{x}) = f(\overline{x}) - \langle v,\overline{x}\rangle$. Employing (11) again gives us

$$f(x) - \langle v,x\rangle + \frac{1}{2\mu}\|x-y\|^2 \geq (f - \langle v,\cdot\rangle)_\mu(\overline{x}) + \frac{1}{2\mu}\|x-y\|^2 - \frac{r}{2}\|x-\overline{x}\|^2$$

for any x and $y \in H$, which in turn yields by taking the infimum over $x \in H$ that

$$\begin{aligned}(f - \langle v,\cdot\rangle)_\mu(y) &\geq (f - \langle v,\cdot\rangle)_\mu(\overline{x}) + \inf_{x\in H}\left(\frac{1}{2\mu}\|x-y\|^2 - \frac{r}{2}\|x-\overline{x}\|^2\right)\\ &= (f - \langle v,\cdot\rangle)_\mu(\overline{x}) - \frac{r}{2(1-r\mu)}\|\overline{x}-y\|^2.\end{aligned}$$

The latter ensures that $0 \in \partial_{\mathrm{p}}(f - \langle v,\cdot\rangle)_\mu(\overline{x})$. Applying further (8), we get

$$0 \in \partial_{\mathrm{p}}\left(f - \langle v,\cdot\rangle\right)_\mu(\overline{x}) = \partial_{\mathrm{p}} f_\mu(\overline{x}+\mu v) - v,$$

which is equivalent to $v \in \partial_{\mathrm{p}} f_\mu(\overline{x}+\mu p)$ and so completes the proof. □

Remark 3.9. If we define $\overline{r}(f,\overline{x},v)$ as the smallest $r > 0$ for which (11) holds, then clearly $\overline{r}(f,\overline{x},0) = \overline{r}(f,\overline{x})$. It can be shown furthermore that $\overline{r}(f,\overline{x},v)$ is positive for any $\overline{x} \in \operatorname{dom} f$ such that $v \in \partial_{\mathrm{p}} f(\overline{x}) \neq \emptyset$.

To conclude this section, we establish an important representation of the basic subdifferential for a prox-bounded function via proximal subgradients of the infimal convolution (1) that is useful in deriving the main results below.

Theorem 3.10. *Assume that* $f : H \to (-\infty, \infty]$ *is lsc and prox-bounded at* $x \in \operatorname{dom} f$. *Then we have the representation*

$$(12)\quad \partial f(x) = \operatorname*{Lim\,sup}_{m\to\infty} \left\{ \partial_{\rm p} f_{\lambda_m}(x_m) \big|\ x_m \to x,\ f_{\lambda_m}(x_m) \to f(x),\ \lambda_m \downarrow 0 \right\}.$$

Proof. To justify the inclusion "$\subset$" in (12), take $v \in \partial f(x)$ and, by definition of the basic subdifferential, find sequences $x_m \to^f x$ and $v_m \xrightarrow{w} v$ as $m \to \infty$ such that $v_m \in \partial_{\rm p} f(x_m)$ for all $m \in I\!N$. Having $\overline{r}(f, x_m, v_m) > 0$ by Remark 3.9, we select a sequence $\lambda_m \in (0, 1/\overline{r}(f, x_m, v_m))$ with $\lambda_m \downarrow 0$ as $m \to \infty$. Then Proposition 3.8 ensures that $v_m \in \partial_{\rm p} f_{\lambda_m}(x_m + \lambda_m v_m)$ and

$$(13)\qquad f(x_m) + \frac{\lambda_m}{2}\|v_m\|^2 = f_{\lambda_m}(x_m + \lambda_m v_m) \ \text{ for all } \ m \in I\!N.$$

Taking into account the above selection of the sequence $\lambda_m \downarrow 0$ and $x_m \to^f x$ as well as the boundedness of the set $\{v_m|\ m \in I\!N\}$ in H due to the sequential weak convergence of $v_m \xrightarrow{w} v$, we get from (13) that

$$f_{\lambda_m}(x_m + \lambda_m v_m) \to f(x) \ \text{ as } \ m \to \infty.$$

Denoting $z_m := x_m + \lambda_m v_m$ allows us to represent v as the weak limit of the proximal subgradients $v_m \in \partial_{\rm p} f_{\lambda_m}(z_m)$ with $z_m \to x$, $\lambda_m \downarrow 0$, and $f_{\lambda_m}(z_m) \to f(x)$ as $m \to \infty$, which justifies the inclusion "$\subset$" in (12).

To prove the opposite inclusion "$\supset$" in (12), fix any

$$v \in \operatorname{Lim\,sup} \left\{ \partial_{\rm p} f_{\lambda_m}(x_m) \big|\ x_m \to x,\ f_{\lambda_m}(x_m) \to f(x),\ \lambda_m \downarrow 0 \right\}$$

and find, by definition of the Painlevé-Kuratowski outer limit in Section 1, sequences of $(\lambda_m, x_m, v_m) \in \mathbb{R} \times H \times H$ such that $v_m \in \partial_{\rm p} f_{\lambda_m}(x_m)$ with the convergences $\lambda_m \downarrow 0$, $x_m \to x$, $f_{\lambda_m}(x_m) \to f(x)$, and $v_m \xrightarrow{w} v$ as $m \to \infty$. It follows now from Lemma 3.4 that, for all $m \in I\!N$, we have

$$v_m \in \partial_{\rm p} f(x_m - \lambda_m v_m) \quad \text{and} \quad f(x_m - \lambda_m v_m) = f_\lambda(x_m) - \frac{\lambda_m}{2}\|v_m\|^2.$$

Denoting $z_m := x_m - \lambda_m v_m$ and using the arguments similar to those in the proof of the inclusion "$\subset$" above, we conclude that

$$z_m \to x \ \text{ and } \ f(z_m) \to f(x) \ \text{ as } \ m \to \infty$$

with $v_m \in \partial_{\rm p} f(z_m)$ and $v_m \xrightarrow{w} v$. Thus $v \in \operatorname{Lim\,sup}_{z \to^f x} \partial_{\rm p} f(z) = \partial f(x)$, which justifies the inclusion "$\supset$" in (12) and completes the proof of the theorem. □

Remark 3.11. The need for using *weak* closure in the construction of the basic subdifferential from Definition 3.1(ii), and hence in Theorem 3.10, is highlighted by the example $f(x) := -d_C(x)$ with the negative distance function in the classical Hilbert space ℓ_2, where C is the norm-compact Hilbert cube

$$C := \left\{ x \in \ell_2 \big|\ \|x_n\| \le 1/2^n, n = 1, 2, \ldots \right\}.$$

Since f is concave and (globally) Lipschitz, every proximal subgradient at $x \notin C$ is in fact a Fréchet derivative and necessarily has norm one; see, e.g., [11, Theorem 5.3.4]. Moreover, C is norm-compact and so has empty interior. Hence $\partial_{\rm p} f(x) = \emptyset$ for all $x \in C$. It follows that while $\partial f(0) = \{0\}$, since C is symmetric and densely spanning, the set of norm-cluster points of nearby proximal normals is empty. Note finally that we may represent the above function f explicitly as

$$f(x) = -\sqrt{\sum_{n\geq 1} (\max\{0, (|x_n| - 1/2^n)\}^2)}.$$

4. Infimal convolutions of prox-regular functions

This section is devoted to the further analysis of infimal convolutions applied to prox-regular functions and their modifications in Hilbert space. Recall first the basic definitions taken from [27, 30].

Definition 4.1. *Let $f : H \to (-\infty, \infty]$, and let $\overline{x} \in \operatorname{dom} f$. We say that f is* PROX-REGULAR *at $\overline{x}$ for $\overline{v} \in \partial f(\overline{x})$ if there exist $\varepsilon > 0$ and $r \geq 0$ such that*

$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2 \qquad \text{for all } x' \in B_\varepsilon(\overline{x}) \tag{14}$$

whenever $x \in B_\varepsilon(\overline{x})$, $|f(x) - f(\overline{x})| < \varepsilon$, and $v \in \partial f(x)$ with $\|v - \overline{v}\| < \varepsilon$. If this holds for every $\overline{v} \in \partial f(\overline{x})$, we say that f is prox-regular at $\overline{x}$.

Definition 4.2. *A function $f : H \to (-\infty, \infty]$ is* SUBDIFFERENTIALLY CONTINUOUS *at $\overline{x} \in \operatorname{dom} f$ for $\overline{v} \in \partial f(\overline{x})$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(\overline{x})| \leq \varepsilon$ whenever $|x - \overline{x}| \leq \delta$ and $|v - \overline{v}| \leq \delta$ with some $v \in \partial f(x)$. If this occurs for all $\overline{v} \in \partial f(\overline{x})$, we say that f is subdifferentially continuous at $\overline{x}$.*

The following proposition provides a useful estimate of the prox-boundedness modulus $r(f, x, v)$ defined in Remark 3.9 in the case of prox-regular functions and also under the additional subdifferential continuity requirement.

Proposition 4.3. *Let $f : H \to (-\infty, \infty]$ be prox-regular at $\overline{x} \in H$ for $\overline{v} \in \partial f(\overline{x})$ with some constants $\varepsilon > 0$ and $r > 0$, and let also f be prox-bounded. Then there is $\eta > 0$ such that $\overline{r}(f, x, v) \leq \eta$ for all $\|x - \overline{x}\| < \varepsilon/2$ with $|f(x) - f(\overline{x})| < \varepsilon$ and all $\|v - \overline{v}\| < \varepsilon$ with $v \in \partial f(x)$. If in addition f is subdifferentially continuous at $\overline{x}$ for $\overline{v}$, then we may drop the condition $|f(x) - f(\overline{x})| < \varepsilon$ above, perhaps after some reducing the value of $\varepsilon > 0$.*

Proof. By the assumed prox-regularity of f at $\overline{x}$ for $\overline{v}$, find the corresponding positive constants ε and r. Let $x \in B_\varepsilon(\overline{x})$ be such that $|f(x) - f(x')| < \varepsilon$, and let $v \in \partial f(x)$ be such that $\|v - \overline{v}\| < \varepsilon$. Take $x' \in B_\varepsilon(x)$ and, by the underlying prox-regularity inequality, get

$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2. \tag{15}$$

Using further the prox-boundedness of f, find $\alpha \in \mathbb{R}$ and $\gamma > 0$ for which

$$f(z) \geq \alpha - \frac{\gamma}{2}\|z\|^2 \quad \text{whenever} \quad z \in H.$$

Let us now justify the existence of a positive number r_1 such that for all $x \in B_\varepsilon(\overline{x})$ with $|f(x) - f(x')| < \varepsilon$ and for all $v \in \partial f(x)$ with $\|v - \overline{v}\| < \varepsilon$ we have

$$f(x') \geq \alpha - \frac{\gamma}{2}\|x'\|^2 \geq f(x) - \langle v, x' - x\rangle - \frac{r_1}{2}\|x' - x\|^2 \tag{16}$$

whenever $x' \notin B_\varepsilon(x)$. Indeed, choose $r_1 > 0$ so that the inequality

$$\frac{r_1 - \gamma}{2}\varepsilon \geq \frac{1}{\varepsilon}\Big(f(x) - \alpha + \frac{\gamma}{2}\|x\|^2\Big) + \|\gamma x + v\|$$

holds for any $x \in B_\varepsilon(\overline{x})$ with $|f(x) - f(x')| < \varepsilon$ and any $v \in \partial f(x)$ with $\|v - \overline{v}\| < \varepsilon$. Then taking $x' \notin B_\varepsilon(x)$, we get the estimate

$$\frac{r_1 - \gamma}{2}\|x' - x\| \geq \frac{1}{\|x' - x\|}\Big(f(x) - \alpha + \frac{\gamma}{2}\|x\|^2\Big) + \Big\langle v + \gamma x, \frac{x' - x}{\|x' - x\|}\Big\rangle,$$

which after simplification yields (16). Combining (15) and (16), we can see that the number $\eta := \max\{r, r_1\}$ is the one ensuring the conclusion of the proposition in the case of prox-regular and prox-bounded functions. Finally, the freedom to drop the condition $|f(x) - f(\overline{x})| < \varepsilon$ in the proposition for subdifferentially continuous functions follows directly from the definition of subdifferential continuity. □

A concept introduced in [7] is also relevant here.

Definition 4.4. *A function $f : H \to (-\infty, \infty]$ is* UNIFORMLY PROX-REGULAR *on a set $E \subset H$ if there are $\varepsilon > 0$ and $r > 0$ such that for any $\overline{x} \in E$ and $\overline{v} \in \partial f(\overline{x})$ we have*

$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x' - x\|^2 \qquad \text{for all } x' \in B_\varepsilon(\overline{x})$$

whenever $v \in \partial f(x)$ with $\|v - \overline{v}\| < \varepsilon$ and $\|x - \overline{x}\| < \varepsilon$ with $|f(x) - f(\overline{x})| < \varepsilon$. We say that f is LOCALLY UNIFORMLY PROX-REGULAR *around x_0 if E can be taken as a neighborhood of x_0, i.e., $E = B_\delta(x_0)$ for some $\delta > 0$.*

The next result clarifying the definition of local uniform prox-regularity is taken from [7, Proposition 3.3].

Proposition 4.5. *A function $f : H \to (-\infty, \infty]$ is uniformly prox-regular around $x_0 \in H$ if and only if there are some $\varepsilon > 0$ and $r > 0$ such that for any $x, x' \in B_\varepsilon(x_0)$ and $v \in \partial f(x)$ we have the estimate*

$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{r}{2}\|x - x'\|^2.$$

We intend to show that the *infimal convolution* of a *prox-regular* function is actually *locally uniformly prox-regular*. Observe that the proof given below does not rely on the usual path to deal with infimal convolutions of prox-regular functions via certain f-attentive subdifferential localizations developed, e.g., in [27] and [8]. As a consequence of our major result given in Theorem 4.9, we establish the $C^{1,1}$ *property* for infimal convolutions of prox-regular functions, which will be used in Section 5 to derive the desired property of subdifferentials of prox-regular functions announced in Section 1.

To proceed in this direction, let us first present some relatively elementary observations regarding infimal convolutions of arbitrary lsc functions in Hilbert spaces needed in the sequel. We impose the following assumptions on the function $f : H \to (-\infty, \infty]$ under consideration:

$$\text{(17)} \qquad \begin{cases} f \text{ is lsc around } 0,\ f(0) = 0, \text{ and there exists} \\ \vartheta > 0 \text{ such that } f(x) > -\frac{\vartheta}{2}\|x\|^2 \ \text{ for all } \ x \neq 0. \end{cases}$$

This easily implies that $f_\lambda(0) = 0$ and $P_\lambda(0) = \{0\}$ for (1) and (2), respectively, when $0 < \lambda < 1/\vartheta$. Observe that the assumptions made can always be enforced via an appropriate translation of the graph of f.

We begin with some estimates that depend only on assumptions (17) and do not yet call for f to be prox-regular. The following result was established in [27, Lemma

4.1] in finite dimensions. The proof given therein holds with no actual change in the Hilbert space settings, and thus it is omitted here.

Lemma 4.6. *Assuming* (17)*, take any* $\lambda \in (0, 1/\vartheta)$*,* $\rho \geq 0$*, and* $x, x' \in H$*. If*

$$f(x') + \frac{1}{2\lambda}\|x' - x\|^2 \leq f_\lambda(x) + \rho,$$

then we have the estimates

$$\begin{aligned}\|x'\| &\leq 2\alpha\|x\| + \sqrt{2\lambda\alpha\rho},\\ f(x') &\leq \frac{1}{2\lambda}\|x\|^2 + \rho,\\ f(x') &\geq -\frac{\vartheta}{2}(2\alpha\|x\| + \sqrt{2\lambda\alpha\rho})^2,\end{aligned}$$

where the number α *is defined by* $\alpha := (1 - \lambda\vartheta)^{-1}$*.*

The next result concerns Lipschitzian behavior of infimal convolutions; parts (i) and (iii) can be found in [27, Proposition 4.2] for finite-dimensional spaces.

Proposition 4.7. *Assuming* (17) *and taking any* $\lambda \in (0, 1/\vartheta)$ *and* $L > 0$*, there is* $\delta > 0$ *such that the following conditions hold:*

(i) $\|x'\| < L$*,* $|f(x')| < L$*, and* $\|x - x'\| < \lambda L$ *for all* $x \in B_\delta(0)$ *and* $x' \in P_\lambda(x)$*.*

(ii) *For any* $x \in B_\delta(0)$ *we have the representation*

$$f_\lambda(x) = \inf_{x' \in B_L(0)} \left(f(x') + \frac{1}{2\lambda}\|x - x'\|^2 \right).$$

(iii) f_λ *is Lipschitz continuous on* $B_\delta(0)$ *with modulus* L*.*

Proof. Pick $\lambda \in (0, 1/\vartheta)$ and let $\alpha := (1 - \lambda\vartheta)^{-1}$. Choose $\beta, \delta > 0$ such that

$$2\alpha\delta + \sqrt{2\lambda\alpha\beta} < L,\ \frac{1}{2\lambda}\delta^2 < L,\ (1 + 2\alpha)\delta + \sqrt{2\lambda\alpha\beta} < \lambda L,\ \frac{\vartheta}{2}(2\alpha\delta)^2 < L.$$

Now take $x \in B_\delta(0)$ and start with proving (i). Given $x' \in P_\lambda(x)$, invoke Lemma 4.6 with $\rho = 0$ to obtain the estimates

$$\|x'\| \leq 2\alpha\delta < L,\ f(x') \leq \delta^2/(2\lambda) < L,\ -f(x') \leq 2\vartheta\alpha^2\delta^2 < L,$$

$$\|x - x'\| \leq \|x\| + \|x'\| \leq \delta + 2\alpha\delta < \lambda L,$$

which surely justify all the properties in (i).

To prove (ii), suppose that some $x' \in H$ satisfies the inequality

$$f(x') + \frac{1}{2\lambda}\|x' - x\|^2 \leq f_\lambda(x) + \beta.$$

Then by Lemma 4.6 we have

$$\|x'\| \leq 2\alpha\delta + \sqrt{2\lambda\alpha\beta} < L,$$

which clearly justifies the representation in (ii). To prove finally (iii), observe that in (ii) we take the infimum of the functions

$$\Phi_{x'} : x \mapsto f(x') + \frac{1}{2\lambda}\|x - x'\|^2 \tag{18}$$

over $x' \in B_L(0)$. These functions are differentiable with their gradients computed by $\nabla\Phi_{x'}(x) = \langle x - x', \cdot\rangle/\lambda$. Hence we have

$$\|\nabla\Phi_{x'}(x)\| = \frac{\|x - x'\|}{\lambda} < \frac{\lambda L}{\lambda} = L.$$

Consequently, functions (18) are Lipschitz continuous on $B_\delta(0)$ with modulus L, and thus the infimal convolution f_λ must have the same property. This justifies (iii) and completes the proof of the proposition. □

The following result is taken from [27, Proposition 4.3], where it is formulated in finite dimensions while the proof given applies to the Hilbert space setting, and so it is omitted here. Recall from Remark 3.5 that $P_\lambda(x) \neq \emptyset$ whenever $\partial_{\rm p} f_\lambda(x) \neq \emptyset$.

Proposition 4.8. *Assuming* (17) *and taking any* $\lambda \in (0, 1/\vartheta)$, *there exists* $\delta > 0$ *such that for all* $x \in B_\delta(0)$ *we have the relationships:*

(i) $\partial f_\lambda(x) \subset \big\{\lambda^{-1}(x - x') : x' \in P_\lambda(x)\big\}$;
(ii) $x' \in P_\lambda(x)$ *implies* $\lambda^{-1}(x - x') \in \partial f(x')$, *i.e.,* $x' \in (I + \lambda\partial f)^{-1}(x)$.

Now we are in a position to establish the uniform prox-regularity of f_λ. It will be shown first that a prox-regular function satisfying (17) has a uniform prox-regular infimal convolution, and then the assumptions of (17) will be removed.

Theorem 4.9. *Let* $f : H \to (-\infty, \infty]$ *be prox-regular at* $\overline{x} = 0$ *for* $\overline{p} = 0$ *with constant* $r := \vartheta > 0$ *in* (14), *and let the assumptions in* (17) *be satisfied. Then, for any* $\lambda \in (0, 1/\vartheta)$, *the infimal convolution* f_λ *is locally uniformly prox-regular at* $\overline{x} = 0$ *with respect to* $\frac{\vartheta}{1-\lambda\vartheta}$. *In particular, the latter implies that* f_λ *is para-convex and* $C^{1,1}$ *in some neighborhood of* $\overline{x} = 0$.

Proof. Take $\lambda \in (0, 1/\vartheta)$. Suppose that $\delta > 0$ has all the properties from both Propositions 4.7 and 4.8 with some $L > 0$. Take $x \in B_\delta(0)$, and an arbitrary subgradient $p \in \partial f_\lambda(x)$. Let further $x' \in P_\lambda(x)$ be such that $p = \lambda^{-1}(x - x')$; see Proposition 4.8. Note that $p \in \partial f(x')$, and thus by Proposition 4.7 we have the estimates $\|x' - \overline{x}\| = \|x'\| < L$, $|f(x') - f(\overline{x})| = |f(x')| < L$ and $\|p\| = \|p - \overline{p}\| = \lambda^{-1}\|x - x'\| < L$.

Invoking next the definition of prox-regularity of f at $\overline{x} = 0$ for $\overline{p} = 0$, we get

$$f(z) - f(x') \geq \langle p, z - x'\rangle - \frac{\vartheta}{2}\|z - x'\|^2 \quad \text{for all} \quad z \in B_L(0),$$

which implies, for any $y \in H$, that

$$\begin{aligned} f(z) + \frac{1}{2\lambda}\|z - y\|^2 - \left(f(x') + \frac{1}{2\lambda}\|x - x'\|^2\right) \\ \geq \frac{1}{2\lambda}\left(\|z - y\|^2 - \|x - x'\|^2\right) + \langle p, z - x'\rangle - \frac{\vartheta}{2}\|z - x'\|^2. \end{aligned} \tag{19}$$

Since $x' \in P_\lambda(x)$, we have $f(x') + \frac{1}{2\lambda}\|x - x'\|^2 = f_\lambda(x)$. Furthermore, Proposition 4.7(ii) ensures the representation

$$f_\lambda(y) = \inf_{z \in B_L(0)} \left(f(z) + \frac{1}{2\lambda}\|z - y\|^2\right)$$

for all $y \in B_\delta(0)$. Since the function $z \mapsto f(z) + \frac{1}{2\lambda}\|z - y\|^2$ is minorized by a positive definite quadratic in z, by employing (19) we can find a minorant of $f_\lambda(y) - f_\lambda(x)$

computing the minimal point of this quadratic. Differentiating the right hand side of (19) with respect to z and using $p = \lambda^{-1}(x - x')$, we arrive at the equation

$$0 = \frac{1}{\lambda}(z - y) + \frac{1}{\lambda}(x - x') - \vartheta(z - x'),$$

which has the exact solution

$$z = x' - \frac{1}{1 - \lambda\vartheta}(x - y).$$

Then the elementary transformations give us the following:

$$\begin{aligned}
f_\lambda(y) - f_\lambda(x) \geq{}& \frac{1}{2\lambda}\left(\left\|x' - y - \frac{1}{1 - \lambda\vartheta}(x - y)\right\|^2 - \|x - x'\|^2\right) \\
&+ \left\langle p, -\frac{1}{1 - \lambda\vartheta}(x - y)\right\rangle - \frac{\vartheta}{2}\left\|\frac{1}{1 - \lambda\vartheta}(x - y)\right\|^2 \\
&= \frac{1}{2\lambda}\left(\left\|(x' - x) + \left(1 - \frac{1}{1 - \lambda\vartheta}\right)(x - y)\right\|^2 - \|x - x'\|^2\right) \\
&+ \frac{1}{1 - \lambda\vartheta}\langle p, y - x\rangle - \frac{\vartheta}{2}\left(\frac{1}{1 - \lambda\vartheta}\right)^2\|x - y\|^2 \\
&= \frac{1}{2\lambda}\left(-\frac{2\lambda\vartheta}{1 - \lambda\vartheta}\langle x' - x, x - y\rangle + \left(\frac{\lambda\vartheta}{1 - \lambda\vartheta}\right)^2\|x - y\|^2\right) \\
&+ \frac{1}{1 - \lambda\vartheta}\langle p, y - x\rangle - \frac{\vartheta}{2(1 - \lambda\vartheta)^2}\|x - y\|^2 \\
&= \left(\frac{1}{1 - \lambda\vartheta} - \frac{\lambda\vartheta}{1 - \lambda\vartheta}\right)\langle p, y - x\rangle \\
&- \left(\frac{\vartheta}{2(1 - \lambda\vartheta)^2} - \frac{1}{2\lambda}\left(\frac{\lambda\vartheta}{1 - \lambda\vartheta}\right)^2\right)\|x - y\|^2 \\
&= \langle p, y - x\rangle - \frac{\vartheta}{2(1 - \lambda\vartheta)}\|x - y\|^2 \qquad \text{for all } x, y \in B_\delta(0).
\end{aligned}$$

Since the subgradient $p \in \partial f_\lambda(x)$ was chosen arbitrary, we get precisely the local uniform prox-regularity of f_λ at $\overline{x} = 0$ with respect to $\frac{\vartheta}{1-\lambda\vartheta}$. Then the para-convexity of f_λ follows from [7, Proposition 3.6e]. Taking finally into account that the infimal convolution is always para-concave, we deduce from Theorem 2.3 that f_λ is actually $C^{1,1}$ around $\overline{x} = 0$ and thus complete the proof of the theorem. □

The next result, which is a consequence of Theorem 4.9, shows that the additional (to prox-regularity) assumptions of Theorem 4.9 can be removed. It is easy to observe this by various translations regarding $p = 0$, $\overline{x} = 0$, and $f(\overline{x}) = 0$. To remove all the assumptions in (17), we employ a rather standard trick that reveals how benign the prox-boundedness assumption is while considering only local properties of lower semicontinuous prox-regular functions.

Corollary 4.10. *Let $f : H \to (-\infty, \infty]$ be lsc and prox-regular at $\overline{x} \in \operatorname{dom} f$ for $\overline{p} \in \partial f(\overline{x})$ with respect to $\vartheta > 0$. Then, for any $\lambda \in (0, 1/\vartheta)$, the function*

$$x \mapsto f_\lambda(x + \lambda\overline{p})$$

is locally uniformly prox-regular at $\overline{x}$ *for* $\overline{p}$ *with respect to* $\frac{\vartheta}{1-\lambda\vartheta}$. *In particular, the infimal convolution* f_λ *is para-convex inside some neighborhood of* $\overline{x}+\lambda\overline{p}$.

Proof. Observe that the lower semicontinuity of f around $\overline{x}$ ensures the existence of a neighborhood $B_\delta(\overline{x})$ on which f is bounded from below. To remove the assumptions of $\overline{x}=0$ and $f(\overline{x})=0$, consider the following translations. Apply first Theorem 4.9 to the function

$$\widetilde{f}(x) := f(x+\overline{x}) + \delta_{B_\delta(\overline{x})}(x+\overline{x}) - f(\overline{x}).$$

Assuming first that $\overline{p}=0$, we have by the prox-regularity of f at $\overline{x}$ for $\overline{p}=0$ that

$$f(x) \geq f(\overline{x}) - \frac{r}{2}\|x-\overline{x}\|^2 \quad \text{implying} \quad \widetilde{f}(x) \geq -\frac{r}{2}\|x\|^2$$

for all $x \in B_\delta(\overline{x})$. The local properties of f around $\overline{x}$ are not affected by either the localization to $B_\delta(\overline{x})$ or by the translation. Apply then Theorem 4.9 to $\widetilde{f}$ arriving in this way at the desired result for f at $\overline{x}$ with the only assumption that $\overline{p}=0$.

To remove the latter assumption, we perform a translation $\widehat{f} := f - \langle \overline{p}, \cdot\rangle$ so that $0 \in \partial\left(f - \langle \overline{p}, \cdot\rangle\right)(\overline{x})$. Deduce then that $\left(f - \langle \overline{p}, \cdot\rangle\right)_\lambda$ is prox-regular at $\overline{x}$ for $p=0$, which implies by Lemma 3.6 that the same holds for the function

$$\left(f - \langle \overline{p}, \cdot\rangle\right)_\lambda(x) = f_\lambda(x+\lambda\overline{p}) - \langle \overline{p}, x\rangle - \frac{\lambda}{2}\|\overline{p}\|^2.$$

Applying finally to the above function $\widehat{f}$ the elementary subdifferential sum rule from [25, Proposition 1.107(ii)], we conclude that the function $x \mapsto f_\lambda(x+\lambda\overline{p})$ is prox-regular at $\overline{x}$ for $\overline{p}$, which completes the proof of the corollary. □

Combining the developments presented above, we arrive at the following important conclusion; cf. [6, 8, 19, 27] for related results in finite-dimensional and infinite-dimensional settings. Note that all the previous considerations in the literature are based on Minty's theory of maximal monotone operators, while we develop a significantly different geometric approach to the $C^{1,1}$ property of infimal convolutions.

Theorem 4.11. *Let* $f : H \to (-\infty, \infty]$ *be lsc and prox-regular at* $\overline{x}$ *for* $\overline{p}$ *with constant* $r := \vartheta$ *in* (14). *Then, for any* $\lambda \in (0, 1/\vartheta)$, *the infimal convolution* f_λ *is a* $C^{1,1}$ *function throughout some neighborhood of* $\overline{x}+\lambda\overline{p}$.

Proof. As mentioned above, the infimal convolution is always a para-concave function. Its para-concavity in Hilbert spaces is established in Theorem 4.9 and Corollary 4.10. Applying finally the characterization of Theorem 2.3, we conclude that f_λ is $C^{1,1}$ around $\overline{x}+\lambda\overline{p}$ and thus complete the proof of the theorem. □

Note that the neighborhood in Theorem 4.11 does depend on λ. We show in the next section (Lemma 5.3) that under some stronger assumptions it is possible to select such a neighborhood *uniformly* with respect to all λ sufficiently small.

5. Lipschitzian Properties of Subdifferentials

The final section of this paper is devoted to applications of the results obtained above to the study of Lipschitzian properties of subdifferential mappings for prox-regular and subdifferentially continuous functions in Hilbert spaces, which was actually the main original motivation for this research. We intend to show that natural extensions of local Lipschitz continuity to set-valued mappings implies, for the case of *subdifferential mappings* generated by *prox-regular* and subdifferentially continuous functions, that the subdifferential mapping is in fact locally *single-valued* and hence the function in question is *locally* $C^{1,1}$.

Properties of this type have been well recognized for subdifferentials of *convex* functions due to their *monotonicity*. This essentially goes back to Kenderov [21] who was the first to observe that the monotonicity and semi/inner continuity of a set-valued mapping implied its local single-valuedness in general infinite-dimensional frameworks. More recently, Levy and Poliquin [23] have extended Kenderov's result, in the case of finite-dimensional spaces, to some generalized notions of monotonicity. Furthermore, they applied it to appropriate Lipschitzian properties of set-valued mapping and applied to subdifferential mappings generated by prox-regular and subdifferentially continuous functions in finite dimensions.

The main result of this section, Theorem 5.4, is an extension of [23, Theorem 3.1] to the case of *Hilbert spaces*. Note that, in contrast to the heavily finite-dimensional technique of [23] involving generalized monotonicity, our approach based on *infimal convolutions* is completely different from that in [23] and allows us to proceed in the general Hilbert space setting.

It is worth also mentioning that the possibility to reduce a set-valued Lipschitzian behavior to a locally *single-value* one plays a key role in many aspects of optimization and variational analysis; in particular, in stability and sensitivity issues related to Robinson's *strong regularity* [29] of solutions maps to parametric generalized equations and variational inequalities. In this paper we are not going further these directions referring the reader to [4, 11, 23, 25, 29, 30] and the bibliographies therein. See, however, some related discussions in Remark 5.5.

Given a set-valued mapping $F : H \rightrightarrows H$, recall that it is *Lipschitz-like* (or has the Aubin property) around $(x, z) \in \operatorname{Graph} F$ with modulus $L \geq 0$ if there exist constants $\delta > 0$ and $\varepsilon > 0$ such that

$$F(x') \cap B_\delta(z) \subset F(x'') + L\|x' - x''\|\overline{B}_1(0) \tag{20}$$

for all $x', x'' \in B_\varepsilon(x)$. This property was introduced in [3] under the name of the "pseudo-Lipschitz property" of F at (x, z). The latter terminology in fact is not really appropriate to describe the essence of (20), since a common meaning is "false" while (20) turns out to be the most natural extension of the classical local Lipschitz continuity to set-valued mappings. It reduces to the classical Lipschitz property for single-valued mappings being also a graphical localization of the local Lipschitz continuity of F around $\overline{x}$ in the Hausdorff sense that corresponds to (20) with $\delta = \infty$; see [25, 30] for more discussions.

It has been well recognized that the Lipschitz-like property of F around (x, z) is equivalent to the *metric regularity* of the inverse mapping F^{-1} around (x, z) and also to the *openness at a linear rate* of F^{-1} around this point; see [11, 25, 30] for more details and references. It is worth mentioning that there are complete characterizations of all the above mentioned properties in both finite and infinite dimension settings (including computation of the exact bounds of the corresponding moduli) via the *coderivative* of F, which is counterpart of the basic subdifferential for set-valued mappings; see [24, 26, 30] and the references therein.

The next simple lemma provides convenient descriptions of the Lipschitz-like property of set-valued mappings.

Lemma 5.1. *For $F : H \rightrightarrows H$, $x \in \mathrm{int}(\mathrm{dom}\, F)$, and $z \in F(x)$ the following assertions are equivalent:*

(i) *F is Lipschitz-like around (x, z) with modulus $L \geq 0$.*
(ii) *Given $\kappa > 0$, there exist $\delta > 0$ and $\varepsilon > 0$ such that for all $x', x'' \in B_\varepsilon(x)$ we have the inclusion*

$$F(x') \cap B_\delta(z) \subset F(x'') \cap B_\kappa(z) + L\|x' - x''\|\overline{B_1}(0). \tag{21}$$

(iii) *There exist $\kappa > 0$, $\delta > 0$ and $\varepsilon > 0$ such that for all $x', x'' \in B_\varepsilon(x)$ inclusion* (21) *holds.*

Proof. Let us justify the implication (i) $\Longrightarrow$ (ii). Note that if the Lipschitz-like property in (i) holds for some positive constants δ and $\varepsilon > 0$, then it must also hold for any smaller values of these constants. Choose $\kappa > 0$ and decrease δ and ε if necessary so that $\delta + 2L\varepsilon \leq \kappa$. By the assumed Lipschitz-like property of F around (x, z), for any given $z' \in F(x') \cap B_\delta(z)$ we find $z'' \in F(x'')$ such that

$$\|z' - z''\| \leq L\|x' - x''\| \leq 2L\varepsilon.$$

Since $\|z' - z\| < \delta$, we have $\|z'' - z\| < \kappa$, which proves (ii). The remaining implications (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i) are obvious. □

The next result shows that the Lipschitz-like property around $(0, 0)$ implies a strong form of lower semicontinuity used in what follows.

Lemma 5.2. *Let $F : H \rightrightarrows H$, $x \in \mathrm{int}(\mathrm{dom}\, F)$, and $z \in F(x)$. If F is Lipschitz-like around (x, z) with modulus $L \geq 0$, then it is strongly lower semicontinuous at (x, z) in the sense that for all $\kappa > 0$ there exists $\kappa' > 0$ with*

$$B_{\kappa'}(x) \subset \left\{ v \middle| \; F(v) \cap B_\kappa(z) \neq \emptyset \right\}. \tag{22}$$

Proof. Applying Lemma 5.1(ii), for any $\kappa > 0$ we find $\delta > 0$ and $\kappa' > 0$ such that

$$z \in F(x) \cap B_\delta(z) \subset F(x'') \cap B_\kappa(z) + L\|x - x''\|\overline{B_1}(0)$$

whenever $x'' \in B_{\kappa'}(x)$. This implies, in particular, that

$$F(x'') \cap B_\kappa(z) \neq \emptyset \qquad \text{for all } \; x'' \in B_{\kappa'}(x),$$

which ready yields the strong lower semicontinuity (22). □

Let us now justify that the simultaneous fulfillment of the prox-regularity and subdifferential continuity properties of f and the Lipschitz-like property of ∂f implies the existence of a λ-*independent* neighborhood of the reference point over which Theorem 4.11 holds. The subdifferential continuity helps us to avoid using the f-attentive localization of the function in question. Note that in the case of prox-regularity and subdifferential continuity we have

$$\mathrm{Graph}\, \partial_{\mathrm{p}} f \cap [B_\varepsilon(\overline{x}) \times B_\delta(\overline{v})] = \mathrm{Graph}\, \partial f \cap [B_\varepsilon(\overline{x}) \times B_\delta(\overline{v})] \tag{23}$$

within some ε, δ-neighborhood of $(\overline{x}, \overline{v})$. In the sequel we use an H^2-transformation $T_\lambda : H^2 \to H^2$ defined by

$$T_\lambda(x, z) := (x + \lambda z, z) \;\text{ for all } \lambda > 0 \;\text{ and }\; x, z \in H.$$

Lemma 5.3. *Suppose that $f : H \to (-\infty, \infty]$ is lsc, prox-bounded, and prox-regular at $\overline{x} \in \operatorname{int}(\operatorname{dom}\partial f)$ for $0 \in \partial f(\overline{x})$ with constant $\vartheta > 0$ in (14) and subdifferentially continuous at $\overline{x}$ for $0 \in \partial f(\overline{x})$. Assume furthermore that ∂f has the Lipschitz-like property around $(\overline{x}, 0)$ with modulus $L \geq 0$. Then there are λ_0 and $\varepsilon_0 > 0$ such that for any $\lambda \in (0, \lambda_0)$ the function f_λ is $C^{1,1}$ on $B_{\varepsilon_0}(\overline{x})$ with the Lipschitz constant $\frac{L}{1-\lambda L}$ for its gradient.*

Proof. To simplify notation, suppose with no loss of generality that $\overline{x} = 0$ and repeatedly decrease values of some constants used below instead of introducing new ones. Since f is subdifferentially continuous at 0 for 0 and prox-regular at 0 for $0 \in \partial f(0)$ with constant $\vartheta > 0$, there exist $\varepsilon, \delta > 0$ such that f is prox-regular at any $x \in B_\varepsilon(0)$ for some $z \in B_\delta(0)$ with ϑ, and by (23) we have

$$\operatorname{Graph}\partial f \cap [B_\varepsilon(0) \times B_\delta(0)] = \operatorname{Graph}\partial_{\mathrm{p}} f \cap [B_\varepsilon(0) \times B_\delta(0)] .$$

Apply now Theorem 4.11 at each $x \in B_\varepsilon(0)$ to conclude that for any $\lambda \in (0, \vartheta)$ there is a neighborhood $B_{\varepsilon_\lambda}(x + \lambda z)$ of $x + \lambda z$ on which f_λ is $C^{1,1}$. Take $\kappa \in (0, \delta)$ and find $\kappa' \in (0, \varepsilon)$ according to Lemma 5.2 such that

$$B_{\kappa'}(0) \subset \left\{x' \middle|\, \partial f(x') \cap B_\kappa(0) \neq \emptyset\right\} .$$

Thus for each $x \in B_{\kappa'}(0)$ there is $z \in B_\kappa(0)$ with $(x, z) \in \operatorname{Graph}\partial f$.

It follows from Proposition 3.8 that for all $\lambda \in (0, 1/\vartheta)$ we have

$$T_\lambda\left(\operatorname{Graph}\partial f \cap [B_{\kappa'}(0) \times B_\kappa(0)]\right) \subset \operatorname{Graph}\partial f_\lambda \cap T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right) .$$

Using further Lemma 3.4 together with (23) and the fact that the proximal subdifferential corresponds to the first component of the subjet leads us to the equality

$$T_\lambda\left(\operatorname{Graph}\partial f \cap [B_{\kappa'}(0) \times B_\kappa(0)]\right) = \operatorname{Graph}\partial f_\lambda \cap T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right) .$$

In particular, from Theorem 4.11 and the above considerations we get that the set $\operatorname{Graph}\partial f_\lambda \cap T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right)$ corresponds to the restriction of the graph of a locally Lipschitz function $x \mapsto \nabla f_\lambda(x)$ to a neighborhood $T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right)$ of the origin. Observe next that taking $\overline{\varepsilon} \in (0, \kappa')$ and $\overline{\lambda} > 0$ with $\overline{\varepsilon} + \overline{\lambda}\delta \leq \kappa'$ we arrive at the relationship

$$\|x - \lambda z\| \leq \overline{\varepsilon} + \lambda\delta < \kappa' \ \text{ for all } \ \lambda \in (0, \overline{\lambda}),\ x \in B_{\overline{\varepsilon}}(0), \ \text{ and } \ z \in B_\kappa(0),$$

which can be equivalently written as

$$B_{\overline{\varepsilon}}(0) \times B_\kappa(0) \subset T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right) \ \text{ whenever } \ \lambda \in (0, \overline{\lambda}).$$

As mentioned above, for each $x \in B_{\kappa'}(0)$ there exists $z \in B_\kappa(0)$ such that $(x, z) \in \operatorname{Graph}\partial f$. Hence for all $x \in B_{\overline{\varepsilon}}(0)$ that there is $z \in B_\kappa(0)$ with

$$T_\lambda^{-1}(x, z) \in \operatorname{Graph}\partial f \cap [B_{\kappa'}(0) \times B_\kappa(0)] .$$

The latter allows us to find, for each $x \in B_{\overline{\varepsilon}}(0)$, such $z \in B_\kappa(0)$ that

$$(x, z) \in T_\lambda\left(\operatorname{Graph}\partial f \cap [B_{\kappa'}(0) \times B_\kappa(0)]\right) = \operatorname{Graph}\partial f_\lambda \cap T_\lambda\left(B_{\kappa'}(0) \times B_\kappa(0)\right) .$$

In particular, this implies that f_λ is $C^{1,1}$ on $B_{\overline{\varepsilon}}(0)$.

To complete the proof of the lemma, it remains to justify the choice of the Lipschitz constant for the gradient ∇f_λ in the lemma formulation, by adjusting the corresponding intervals for λ and ε in the discussions above. We proceed by using the assumed Lipschitz-like property of ∂f around $(0, 0)$ with modulus $L > 0$ and employing Lemma 5.1(ii) with our κ chosen above. In this way we find numbers $\delta'' > 0$ and $\varepsilon'' > 0$ such that

$$\partial f(x') \cap B_{\delta''}(0) \subset \partial f(x'') \cap B_\kappa(0) + L\|x' - x''\|\overline{B_1}(0) \tag{24}$$

for all $x', x'' \in B_{\varepsilon''}(0)$. Assume with no loss of generality that $\delta'' < \kappa$.

Then Proposition 4.3 provides us with some $r > 0$ such that $\overline{r}(f, x, v) \le r$ for all elements $x \in B_{\overline{\varepsilon}}(0)$ and $v \in \partial f(x) \cap B_\delta(0)$, after decreasing $\overline{\varepsilon}$ and δ if necessary. Picking $x \in B_{\overline{\varepsilon}}(0)$ and $v \in B_\delta(0)$ and taking into account that the function f is assumed to be prox-bounded, we employ Proposition 3.8 to conclude that the inclusion $v \in \partial f(x)$ implies that $v \in \partial f_\lambda(x + \lambda v)$ for all $\lambda \in (0, 1/r)$. Furthermore, Lemma 3.6 allows us to deduce that the inclusion $v \in \partial f_\lambda(x+\lambda v)$ for such λ yields in turn that $v \in \partial f(x)$. In the remainder of the proof we assume $\overline{\lambda} < 1/r$.

We now show that there exist $\varepsilon_0 \in (0, \overline{\varepsilon})$ and $\lambda_0 \in (0, \overline{\lambda})$ such that for all $\lambda \in (0, \lambda_0)$ the Lipschitz constant of ∇f_λ on $B_{\varepsilon_0}(0)$ can be chosen as $\frac{L}{1-\lambda L}$. Arguing by contradiction, assume that for any arbitrarily small $\varepsilon' > 0$ and $\lambda' > 0$ there are $x', x'' \in B_{\varepsilon'}(0)$ satisfying

$$\|\nabla f_{\lambda'}(x') - \nabla f_{\lambda'}(x'')\| > \frac{L}{1 - \lambda' L}\|x' - x''\| \tag{25}$$

with $\nabla f_{\lambda'}(x') \in B_{\delta''}(0)$ and $\nabla f_{\lambda'}(x'') \in B_\kappa(0)$. Reduce ε' and λ' so that $\varepsilon' + \lambda'\kappa < \varepsilon''$. Setting now

$$x_1 := x' - \lambda'\nabla f_{\lambda'}(x') \in B_{\varepsilon''}(0) \quad \text{and} \quad x_2 := x'' - \lambda'\nabla f_{\lambda'}(x'') \in B_{\varepsilon''}(0)$$

and employing Lemma 3.6, we get the relationships

$$z_1 := \nabla f_{\lambda'}(x') \in \partial f(x_1) \cap B_{\delta''}(0) \quad \text{and} \quad z_2 := \nabla f_{\lambda'}(x'') \in \partial f(x_2) \cap B_\kappa(0).$$

Furthermore, estimate (25) allows us to conclude that

$$\begin{aligned} \|z_1 - z_2\| &> \frac{L}{1 - \lambda' L}\|(x' - \lambda' z_1) - (x'' - \lambda' z_2) + \lambda'(z_1 - z_2)\| \\ &\ge \frac{L}{1 - \lambda' L}\left(\|x_1 - x_2\| - \lambda'\|z_1 - z_2\|\right), \end{aligned}$$

which in turn implies the inequality

$$\left(1 + \frac{\lambda' L}{1 - \lambda' L}\right)\|z_1 - z_2\| > \frac{L}{1 - \lambda' L}\|x_1 - x_2\|$$

and equivalently the estimate

$$\|z_1 - z_2\| > L\|x_1 - x_2\|. \tag{26}$$

Since the subdifferential mapping ∂f is Lipschitz-like around $(0, 0)$ with modulus L and thus satisfies (24), and by the inclusions $x_1 \in B_{\varepsilon''}(0)$, $z_1 \in \partial f(x_1) \cap B_{\delta''}(0)$, and $z_2 \in \partial f(x_2) \cap B_\kappa(0)$ established above, we get from (26) that $x_2 \notin B_{\varepsilon''}(0)$. The latter provides a contradiction, which completes the proof of the lemma. □

Now we are able to prove our principal result showing that the Lipschitz-like property of the subdifferential mapping ∂f for a prox-regular and subdifferentially continuous function $f : H \to (-\infty, \infty]$ implies the continuous differentiability of this function with a locally Lipschitzian derivative.

Theorem 5.4. *Let $f : H \to (-\infty, \infty]$ be lsc, prox-regular, and subdifferentially continuous at $\overline{x} \in \operatorname{int}(\operatorname{dom} \partial f)$ for some $\overline{v} \in \partial f(\overline{x})$. Assume in addition that the subdifferential mapping ∂f is Lipschitz-like with modulus $L \geq 0$ around $(\overline{x}, \overline{v})$. Then there exists $\varepsilon > 0$ such that $\partial f(x) = \{\nabla f(x)\}$ for all $x \in B_\varepsilon(\overline{x})$ with the Lipschitzian derivative $x \mapsto \nabla f(x)$ on $B_\varepsilon(\overline{x})$.*

Proof. Once again translate $\overline{x}$ to 0 and $\overline{v}$ to 0 for convenience.Let $\varepsilon > 0$ be sufficiently small so that f is bounded from below within the neighborhood $\overline{B_\varepsilon}(0)$. Now redefine f to be $f + \delta_{\overline{B_\varepsilon}(0)}$, observing that the assertions of the theorem for the function $f + \delta_{\overline{B_\varepsilon}(0)}$ imply those for the original function f inside a neighborhood of the origin. Clearly $f + \delta_{\overline{B_\varepsilon}(0)}$ is prox-bounded and inherits all the other local properties of f. To simplify notation, we refer to this restriction as to f in what follows and prove the theorem for the latter function.

Take $\lambda_0 > 0$ and further reduce $\varepsilon, \delta > 0$ so that for all $\lambda \in (0, \lambda_0)$ we have that f_λ is $C^{1,1}$ on $B_\varepsilon(0)$ with the Lipschitz constant $L/(1 - \lambda L)$ of the gradient ∇f_λ. This is possible by Lemma 5.3. Then

$$\|\nabla f_\lambda(x) - \nabla f_\lambda(y)\| \leq \frac{L}{1 - \lambda L}\|x - y\|$$

for all $x, y \in B_\varepsilon(0)$. Thus we have

$$\begin{aligned}
-\frac{L}{1 - \lambda L}\|x - y\|^2 &\leq - \|\nabla f_\lambda(x) - \nabla f_\lambda(y)\| \cdot \|x - y\| \\
&\leq \langle \nabla f_\lambda(x) - \nabla f_\lambda(y), x - y\rangle \leq \|\nabla f_\lambda(x) - \nabla f_\lambda(y)\| \cdot \|x - y\| \\
&\leq \frac{L}{1 - \lambda L}\|x - y\|^2 \qquad \text{for all } \; x, y \in B_\varepsilon(0).
\end{aligned}$$

Now use Theorem 3.10 to reconstruct the basic subdifferential of f at x via weak limits of the gradients $\nabla f_{\lambda_m}(x_m)$ of the infimal convolution at the points $x_m \to x$ with $f_{\lambda_m}(x_m) \to f(x)$ and $\lambda_m \downarrow 0$. Taking into account that the bilinear form above is continuous with respect to weak×norm sequential convergence, we get in this way the two inequalities

$$-L\|x - y\|^2 \leq \langle u - v, x - y\rangle \leq L\|x - y\|^2 \tag{27}$$

for all $u \in \partial f(x)$, $v \in \partial f(y)$, and $x, y \in B_\varepsilon(0)$. The left-hand side inequality in (27) implies that the mapping

$$x \mapsto \partial f(x) + Lx = \partial\left(f + \frac{L}{2}\|\cdot\|^2\right)(x)$$

is monotone on $B_\varepsilon(0)$. By Lemma 3.3 we deduce that the function $f + \frac{L}{2}\|\cdot\|^2$ is convex on $B_\varepsilon(0)$. Since it is lsc on the interior of its domain, it is well known to be Lipschitz continuous on $B_\varepsilon(0)$; see, e.g., [11, Theorem 4.1.3]. Hence the right-hand side inequality in (27) implies that the mapping

$$x \mapsto -\,\overline{\text{co}}\,\partial\left(f - \frac{L}{2}\|\cdot\|^2\right)(x) = -\partial_{\text{C}}\left(f - \frac{L}{2}\|\cdot\|^2\right)(x) = \partial_{\text{C}}\left(-f + \frac{L}{2}\|\cdot\|^2\right)(x)$$

is monotone and, by Lemma 3.3, the function $-f + \frac{L}{2}\|\cdot\|^2$ is convex (or $f - \frac{L}{2}\|\cdot\|^2$ is concave) inside $B_\varepsilon(0)$. By Theorem 2.3 we have that $f \in C^{1,1}$, and thus the gradient $\nabla f(x)$ exists for all $x \in B_\varepsilon(0)$. Moreover, the Lipschitz-like property of ∂f with modulus L yields that the mapping $x \mapsto \nabla f(x)$ is Lipschitzian on $B_\varepsilon(0)$ with the same modulus L. This completes the proof of the theorem. □

Remark 5.5. It is worth underlining the importance of the *single-valued subdifferential reduction* of the type given in Theorem 5.4 for the study of *metric regularity* of solution maps to parametric generalized equations in Robinson's form

$$S(x) := \{y \in Y : 0 \in g(x,y) + Q(y)\} \tag{28}$$

with mappings $g : X \times Y \rightrightarrows Z$ and $Q : Y \rightrightarrows Z$ between Banach spaces. As has been well recognized starting with Robinson's seminal contributions (see, in particular, [28, 29]) that model (28) is a convenient form for describing solution maps to parametric variational inequalities, complementarity problems, first-order optimality conditions in parametric optimization, etc.; see, e.g., books [25, 30] with the references and discussions therein. It has been established recently in [1, 4, 18, 26], under various assumptions and with certain modifications, that metric regularity of the solution map S to (28) is equivalent to the Lipschitz-like property of the set-valued field Q in (28). The most interesting cases for applications relate to systems (28) with field mappings Q given in some subdifferential/normal cone forms and their compositions of the types

$$Q(y) = \partial(\varphi \circ h)(y) \quad \text{and} \quad Q(y) = (\partial\varphi \circ h)(y), \tag{29}$$

where $h : Y \to W$ and $\varphi : W \to (-\infty, \infty]$. For such mappings, the aforementioned single-valued subdifferential reduction mandates, under natural assumptions, that the "superpotential" functions φ in (29) exhibit certain smoothness properties that fail to hold for major classes of variational systems. This leads us to conclusions on the *failure of metric regularity* for solution maps to such parametric variational systems; see [26] and also [1, 4, 18] for related results in this direction.

Acknowledgments. The authors are indebted to the referee and Lionel Thibault for many useful remarks that allowed us to significantly improve the original presentation. We also gratefully acknowledge helpful discussions with Warren Hare and René Poliquin on the results presented in the paper.

References

[1] F.H. Aragón Artacho, B.S. Mordukhovich: *Metric regularity and Lipschitzian stability of parametric variational systems.* Nonlinear Anal. 72 (2010), 1149–1170.

[2] H. Attouch, D. Azé: *Approximation and regularization of arbitrary functions on Hilbert spaces by Lasry-Lions method.* Ann. Inst. H. Poincaré 10 (1993), 289–312.

[3] J.-P. Aubin: *Lipschitz behavior of solutions to convex minimization problems.* Math. Oper. Res. 9 (1984), 87–111.

[4] D. Aussel, Y. Garcia, N. Hadjisavvas: *Single-directional property of multivalued maps and variational systems.* SIAM J. Optim. 20 (2009), 1274–1285.

[5] G. Beer: *Topologies on closed and closed convex sets.* Mathematics and its Applications, 268, Kluwer Academic Publ., Dordrecht, 1993.

[6] F. Bernard, L. Thibault: *Prox-regularity of functions and sets in Banach spaces.* Set-Valued Anal. 12 (2004), 25–47.

[7] F. Bernard, L. Thibault: *Uniform prox-regularity of functions and epigraphs in Hilbert spaces.* Nonlinear Anal. 60 (2005), 187–207.

[8] F. Bernard, L. Thibault: *Prox-regular functions in Hilbert spaces.* J. Math. Anal. Appl. 303 (2005), 1–14.

[9] F. Bernard, L. Thibault, N. Zlateva: *Characterizations of prox-regular sets in uniformly convex Banach spaces.* J. Convex Anal. (2006), no. 3–4, 525–559.

[10] J.M. Borwein, J.D. Vanderwerff: *Convex functions: characterizations, constructions and counterexamples.* Cambridge University Press, 2009.

[11] J.M. Borwein, Q.J. Zhu: *Techniques of variational analysis.* CMS Books in Mathematics, 4, Springer, New York, 2005.

[12] F.H. Clarke, Y.S. Ledyaev, R.J. Stern, P.R. Wolenski: *Nonsmooth analysis and control theory.* Graduate Texts in Mathematics, 178, Springer, New York, 1998.

[13] R. Correa, P. Gajardo, L. Thibault: *Subdifferential representation formula and subdifferential criteria for the behavior of nonsmooth functions.* Nonlinear Anal. 65 (2006), 864–891.

[14] M. Crandall, H. Ishii, P.-L. Lions: *User's guide to viscosity solutions of second order partial differential equations.* Bull. Amer. Math. Soc. 27 (1992), 1–67.

[15] A. Eberhard: *Prox-regularity and subjets.* Optimization and Related Topics, Applied Optimization, Ch. 14, pp. 237–313, Appl. Optim., 47, Kluwer Academic Publ., Dordrecht, 2001.

[16] A. Eberhard, M. Nyblom, D. Ralph: *Applying generalized convexity notions to jets.* Generalized convexity, generalized monotonicity: recent results, J.-P. Crouzeix (ed.), pp. 111–157, Nonconvex Optim. Appl., 27, Kluwer Academic Publ., Dordrecht, 1998.

[17] A. Eberhard, R. Sivakumaran, R. Wenczel: *On the variational behaviour of the subhessians of the Lasry-Lions envelope.* J. Convex Anal. (2006), no. 3–4, 647–685.

[18] W. Geremew, B.S. Mordukhovich, N.M Nam: *Coderivative calculus and metric regularity for constraint and variational systems.* Nonlinear Anal. 70 (2009), 529–552.

[19] W.L. Hare, R.A. Poliquin: *Prox-regularity and stability of the proximal mapping.* J. Convex Anal. 14 (2007), 589–606.

[20] J.-B. Hiriart-Urruty, Ph. Plazanet: *Moreau's decomposition theorem revisited.* Analyse non linéaire (Perpignan, 1987). Ann. Inst. H. Poincaré Anal. Non Linéaire 6 (1989), suppl., 325–338.

[21] P. Kenderov: *Semi-continuity of set-valued monotone mappings.* Fund. Math. 88 (1975), 61–69.

[22] J.-M. Lasry, P.-L. Lions: *A remark on regularization in Hilbert spaces.* Israel J. Math. 55 (1986), 257–266.

[23] A.B. Levy, R.A. Poliquin: *Characterizing the single-valuedness of multifunctions.* Set-Valued Anal. 5 (1997), 351–364.

[24] B. S. Mordukhovich: *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions.* Trans. Amer. Math. Soc. 340 (1993), 1–35.

[25] B. S. Mordukhovich: *Variational analysis and generalized differentiation, I: Basic theory, II: Applications.* Fundamental Principles of Mathematical Sciences, 330 and 331, Springer, Berlin, 2006.

[26] B. S. Mordukhovich: *Failure of metric regularity for major classes of variational systems.* Nonlinear Anal. 69 (2008), 918–924.

[27] R.A. Poliquin, R.T. Rockafellar: *Prox-regular functions in variational analysis.* Trans. Amer. Math. Soc. 348 (1996), 1805–1838.

[28] S.M. Robinson: *Generalized equations and their solutions, I: Basic theory.* Math. Program. Study 10 (1979), 128-141.

[29] S.M. Robinson: *Strongly regular generalized equations.* Math. Oper. Res. 5 (1980), 43–62.

[30] R.T. Rockafellar, R. J-B. Wets: *Variational analysis.* Fundamental Principles of Mathematical Sciences, 317, Springer, Berlin, 1998.

[31] W. Schirotzek: *Nonsmooth analysis.* Universitext, Springer, Berlin, 2007.

M. Bačák, School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, Newcastle, New South Wales, Australia

E-mail address: `miroslav.bacak@newcastle.edu.au`

J. M. Borwein, School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, New South Wales, Australia

E-mail address: `jonathan.borwein@newcastle.edu.au`

A. Eberhard, School of Mathematical and Geospatial Sciences, RMIT, GPO Box 2476V, Melbourne, Victoria, Australia 3001.

E-mail address: `andy.eb@rmit.edu.au`

B. S. Mordukhovich, Department of Mathematics, Wayne State University, Detroit, MI 48202, USA.

E-mail address: `boris@math.wayne.edu`