Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 266 (2014) 1185–1237

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Differential properties of the Moreau envelope

Abderrahim Jourani [a], Lionel Thibault [b,*], Dariusz Zagrodny [c]

[a] Université de Bourgogne, Institut de Mathématiques de Bourgogne, UMR 5584 CNRS, BP 47 870, 21078 Dijon Cedex, France

[b] Université de Montpellier II, Institut de Mathématiques et de Modélisation de Montpellier, UMR 5149 CNRS, Case Courrier 051, Place Eugène Bataillon, 34095 Montpellier Cedex, France

[c] Cardinal Stefan Wyszynski University, Faculty of Mathematics and Natural Sciences, College of Science, Dewajtis 5, Warsaw, Poland

Received 12 June 2012; accepted 5 November 2013

Available online 17 December 2013

Communicated by H. Brezis

Abstract

In a vector space endowed with a uniformly Gâteaux differentiable norm, it is proved that the Moreau envelope enjoys many remarkable differential properties and that its subdifferential can be completely described through a certain approximate proximal mapping. This description shows in particular that the Moreau envelope is essentially directionally smooth. New differential properties are derived for the distance function associated with a closed set. Moreover, the analysis, when applied to the investigation of the convexity of Tchebyshev sets, allows us to recover several known results in the literature and to provide some new ones.

Keywords: Moreau envelope; Inf-convolution; Directional differentiability of the Moreau envelope; Differentiability of the distance function; Subregular function; Essentially directionally smooth function; Subdifferential; Tchebyshev sets

1. Introduction

In 1963 J.J. Moreau associated with two extended real-valued functions f, g, defined on a vector space X, the inf-convolution (also called infimal convolution) function $f \,\Box\, g$ defined by

* Corresponding author.

E-mail addresses: jourani@u-bourgogne.fr (A. Jourani), thibault@math.univ-montp2.fr (L. Thibault), d.zagrodny@uksw.edu.pl (D. Zagrodny).

0022-1236/$ – see front matter
http://dx.doi.org/10.1016/j.jfa.2013.11.008

$$f \,\square\, g(x) := \inf_{y\in X}\big[f(y) + g(x-y)\big] \quad \text{for all } x \in X,$$

see [44] for the original definition and for a list of algebraic properties of the inf-convolution operation. In fact, the case of the function $g = \frac{1}{2}\|\cdot\|^2$, where $\|\cdot\|$ is the norm associated with the inner product of a Hilbert space H, had already been considered by Moreau in his earlier note [43] in 1962 dealing with the particular inf-convolution $f \,\square\, \frac{1}{2}\|\cdot\|^2$, that is,

$$x \mapsto \inf_{y\in H}\left[f(y) + \frac{1}{2}\|x-y\|^2\right]. \tag{1.1}$$

In the short note [43] and in the paper [45], Moreau showed that the latter function was of class C^1 whenever the function f was assumed to be proper, convex, and lower semicontinuous on the Hilbert space H. In those papers [43,45], the set-valued mapping (called the "prox" mapping therein and proximal mapping in [51])

$$Pf(x) := \left\{z \in H\colon\ f(z) + \frac{1}{2}\|x-z\|^2 = \left(f \,\square\, \frac{1}{2}\|\cdot\|^2\right)(x)\right\} \quad \text{for all } x \in H,$$

was shown to be single-valued and locally Lipschitz continuous on H, under the same assumptions of properness, convexity, and lower semicontinuity of f. Another particular remarkable property, for f fulfilling the assumptions above, is that (see [45])

$$Pf(x) + Pf^*(x) = x \quad \text{for all } x \in H,$$

where f^* denotes the *Legendre–Fenchel conjugate* of f, that is,

$$f^*(u) := \sup_{y\in H}\big[\langle u, y\rangle - f(y)\big] \quad \text{for all } u \in H;$$

this brought to light an extension to the mapping Pf of an old well-known property of the metric projection on a closed subspace of a Hilbert space.

Some years later, H. Brézis [10,11] revealed the interest for each $\lambda > 0$ to consider the inf-convolution $f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2$, that is, the use of the Hilbert norm $\frac{1}{\sqrt{\lambda}}\|\cdot\|$ in H in place of the initial Hilbert norm $\|\cdot\|$ in (1.1). Indeed, for a proper lower semicontinuous convex function $f : H \to \mathbb{R} \cup \{+\infty\}$ with subdifferential ∂f and the differential inclusion (modelizing constrained parabolic PDE)

$$\begin{cases} \dfrac{du}{dt}(t) + \partial f\big(u(t)\big) \ni 0, \\ u(0) = u_0 \in \mathrm{cl}(\mathrm{dom}\, f), \end{cases} \tag{1.2}$$

Brézis [10, Theorem 1] and [11] showed the interest for u_0 in the *closure* of the effective domain of f to study regularizing properties of the semigroup of the differential inclusion above through the ordinary differential equation

$$\begin{cases} \dfrac{du_\lambda}{dt}(t) + \left(\nabla f \,\square\, \dfrac{1}{2\lambda}\|\cdot\|^2\right)\big(u_\lambda(t)\big) = 0, \\ u_\lambda(0) = u_0 \in \mathrm{cl}(\mathrm{dom}\, f), \end{cases} \tag{1.3}$$

in establishing, via the convergence $f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2(x) \to f(x)$ as $\lambda \downarrow 0$ (see [10, Proposition 10]), that the solution $\{u_\lambda\}_\lambda$ of (1.3) converges to a mapping u which is a solution of (1.2) in a strong sense. H. Attouch [2] and [3, see Theorems 3.24 and 3.66] established in the setting of a reflexive Banach space a general convergence result from which one deduces the Painlevé–Kuratowski approximation of subgradients of a proper lower semicontinuous convex function f by the Fréchet derivative of $f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2$, namely: The graph of $\nabla(f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2)$ Painlevé–Kuratowski converges in $X \times X^*$ endowed with the strong topology to the graph of f as $\lambda \downarrow 0$. Hence in particular any subgradient of f at x is the strong limit of $\nabla(f \,\square\, \frac{1}{2\lambda_n}\|\cdot\|^2)(x_n)$ for some $x_n \to x$ and $\lambda_n \downarrow 0$, that is,

$$\partial f(x) = \limsup_{\lambda\downarrow 0;\, u\to x} \left\{\nabla\left(f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2\right)(u)\right\}. \tag{1.4}$$

We refer to Section 5 for the related definitions and to [3] for several results concerning differential properties of $f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2$ for convex functions f; for potential tools allowing to extend (1.4) to weakly compactly generated Banach spaces we refer to [62,60,61]. Following R.T. Rockafellar and R.J.-B. Wets (see [51]) we will call $f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2$ the *Moreau envelope of f of index λ* and we will denote it as $e_\lambda f$, that is,

$$e_\lambda f(x) := \left(f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2\right)(x).$$

The properties mentioned above show the relevance of the Moreau envelope in Convex Analysis and the need for investigating results on properties of the envelope outside the reflexive Banach space setting.

The Moreau envelope is also relevant for nonconvex nonsmooth functions. The study of the Moreau envelope and proximal mapping of nonconvex functions f started with [9,20,46] in the very particular case when f is convex up to a square, that is, $f = \varphi - c\|\cdot\|^2$ for some real $c \geqslant 0$ and some proper convex lower semicontinuous function $\varphi : H \to \mathbb{R} \cup \{+\infty\}$, and those papers provided applications to Morse theory and trajectories of Hamiltonian systems. Some properties of the Moreau envelope in the more general class of primal lower nice functions are established by L. Thibault and D. Zagrodny in [53] where it is shown: Given a function f from a Hilbert space H into $\mathbb{R} \cup \{+\infty\}$ which is primal lower nice at $\bar{x}$ (see [47,53] for the definition) with $f(\bar{x}) < +\infty$, there exists some $\lambda_0 > 0$ and some closed convex neighborhood U of $\bar{x}$ such that, for $f_U(x) = f(x)$ if $x \in U$ and $f_U(x) = +\infty$ if $x \notin U$, the function $e_\lambda f_U(\cdot) + \frac{1}{2\lambda}\|\cdot\|^2$ is convex on U for any $0 < \lambda < \lambda_0$; this property, combined with the fact that $\frac{1}{2\lambda}\|\cdot\|^2 - e_\lambda f_U(\cdot)$ is always convex on the Hilbert space H, says in particular that $e_\lambda f_U$ is *Fréchet differentiable* on the interior of U since those convexity properties along with the Fréchet differentiability of $\|\cdot\|^2$ ensure that both functions $e_\lambda f_U$ and $-e_\lambda f_U(\cdot)$ are Fréchet subdifferentiable at any point in $\operatorname{int} U$. The convexity property of $e_\lambda f_U(\cdot) + \frac{1}{2\lambda}\|\cdot\|^2$ is used in [53] to show that primal lower nice functions on an open convex set U of H are subdifferentially determined on U, that is, two such functions with the same subdifferential on U are equal up to an additive constant. For another more general class of the so-called prox-regular functions, R.A. Poliquin and R.T. Rockafellar [48] proved a regularity property of the Moreau envelope in establishing the following: For a lower semicontinuous function f on $\mathbb{R}^N$ which is prox-regular at $\bar{x}$ for $\bar{v}$ in the Mordukhovich subdifferential of f at $\bar{x}$ and which is minorized by a quadratic function, there exists $\lambda_0 > 0$ such that for each $0 < \lambda < \lambda_0$ there is a neighborhood U_λ of $\bar{x} + \lambda\bar{v}$ such that $e_\lambda f$ is of class $C^{1,1}$

on U_λ (that is, differentiable on U_λ with the derivative locally Lipschitz continuous on U_λ). This property has been extended to Hilbert spaces in [6] by F. Bernard and L. Thibault. We refer to [48,6] for the definition of prox-regular functions and to Section 2 for other notions. This property of $e_\lambda f$ is exploited on one hand in [48,51] to develop a second order theory for nonsmooth prox-regular functions, and on the other hand in [15,38] to study the differential evolution inclusion (1.2) when the function f is supposed to be merely primal lower nice (instead of being convex).

Recently, important additional properties of the Moreau envelope have been highlighted for any minorized lower semicontinuous function on a Banach space X. Assuming that the norm $\|\cdot\|$ of the Banach space X is locally uniformly convex, R. Cibulka and M. Fabian [12] studied, as an extension of results of S. Dutta [19] related to the distance function d_S from a nonconvex set S, the strong attainment and the differentiability of the Moreau envelope of a certain nonconvex function f at generically many points. If the norm $\|\cdot\|$ is assumed to be uniformly Gâteaux differentiable (instead of being locally uniformly convex) and if $M \subset \operatorname{Dom} Pf$ is dense in a neighborhood of a point $x \in X$, then the paper [12] also established a description of the Clarke subdifferential of $e_1 f$ at x in terms of weak* limits of $D_G(\frac{1}{2}\|\cdot\|^2)(x_n - z_n)$ for $M \ni x_n \to x$ and $z_n \in Pf(x_n)$, extending in this way earlier results of J.M. Borwein, S.P. Fitzpatrick and J.R. Giles [8] for the distance function; where $D_G(\frac{1}{2}\|\cdot\|^2)$ stands for the Gâteaux derivative, see (2.13) below.

Our aim in this paper is threefold. In a first step, under the uniform Gâteaux differentiability of the norm $\|\cdot\|$ we establish a directional subregularity property of the Moreau envelope $e_\lambda f$, see (3.11); we also provide a description of the Clarke subdifferential of $e_\lambda f$ for extended real-valued lower semicontinuous functions f minorized by negative quadratic functions for which the domain of the proximal mapping Pf may be empty, see (3.13). This description allows us to show that the directional subregularity of the Moreau envelope is in fact equivalent to the uniform Gâteaux differentiability of the norm $\|\cdot\|$, and to unify different approaches to the differentiability of the Moreau envelope in this general setup. We also show through this description that the Moreau envelope $e_\lambda f$ is a ∂-essentially smooth function (see the definition in Section 2 and Theorem 3.7) for any subdifferential included in the Clarke subdifferential. In view of results from [54,34] this essential smoothness property says, in particular, that the Moreau envelope belongs to the class of functions determined (or integrable) by the Clarke subdifferential (up to an additive constant). In fact, in several cases, a knowledge on a selection of the subdifferential on some dense subset is sufficient for this determination, that is, we do not need the knowledge on the whole subdifferential but only a single subgradient is enough for this integration property. It is also shown that for every $\lambda > 0$ and every nonconvex weakly closed subset $S \subset X$ the function $-e_\lambda \psi_S$ is ∂_C-eds (see the meaning in Definition 2.1 in the next section) but it is not the Moreau envelope of any function, see Corollary 4.7(e), where ψ_S stands for the indicator function of a set S and ∂_C for the Clarke subdifferential.

In a second step, we apply the results on the Moreau envelope to the investigation of new differential properties of the distance function and we provide new characterizations of convex sets, see for example Theorem 4.3. Using such differential properties of the distance function we derive that weakly closed Tchebyshev subsets of a reflexive Banach space whose norm is uniformly Gâteaux differentiable are convex (see [5,36,55,56] for other results in this line).

As the third step, we establish under the assumption of Asplund property of $(X, \|\cdot\|)$ a description of the Mordukhovich limiting subdifferential $\partial_L f(x)$ in terms of the Fréchet subgradients of the Moreau envelope, providing in this way a partial extension of (1.4) (for other results in this line, we refer to [33,4]).

Differential properties of Moreau upper envelopes on uniformly Gâteaux differentiable Banach spaces will be studied in our forthcoming paper [35].

2. Preliminaries

In this section we collect several facts on generalized derivatives which are used throughout the paper.

Let $(X, \|\cdot\|)$ be a normed vector space. For every nonempty set $S \subset X$ the *distance function* from the set S is denoted by $d_S(\cdot)$ and is defined as

$$d_S(x) = \inf_{u \in S} \|u - x\|, \quad \forall x \in X.$$

The *projection mapping* on S is defined by

$$P_S(x) := \big\{s \in S\colon\ d_S(x) = \|x - s\|\big\}, \quad \forall x \in X. \tag{2.1}$$

For every real $r > 0$ and every $x \in X$ we denote by $B(x, r)$ (resp. $B[x, r]$) the open (resp. closed) ball centered at x and of radius r.

The topological dual space of X is denoted by X^*, its dual norm by $\|\cdot\|_*$, that is,

$$\|x^*\|_* := \sup_{\|u\| \leqslant 1} \langle x^*, u\rangle, \quad \forall x^* \in X^*,$$

where $\langle \cdot,\cdot\rangle$ is the duality pairing between X and X^*. When there is no risk of confusion, we will write $\|x^*\|$ in place of $\|x^*\|_*$. The closed unit ball centered at the origin of X^* (resp. X) is denoted by $\mathbb{B}_{X^*}$ (resp. $\mathbb{B}_X$).

For a function f and a set S, we write $u \xrightarrow{f} x$ and $u \xrightarrow{S} x$ to express $u \to x$ with $f(u) \to f(x)$, and $u \to x$ with $u \in S$, respectively.

Let f be an extended real-valued function on X. The *Mordukhovich limiting subdifferential* of f at x is the set

$$\partial_L f(x) = w^*\text{-}\operatorname*{Lim\,sup}_{u \xrightarrow{f} x,\ \varepsilon \downarrow 0} \partial_{F,\varepsilon} f(u), \tag{2.2}$$

where w^*-Lim sup stands for the weak* *sequential outer* (upper) limit of $\partial_{F,\varepsilon} f(u)$ as $u \to_f x$, that is, the set of all possible w^*-limits $\lim_i u_i^*$ of sequences $\{u_i^*\}_{i \in \mathbb{N}}$ such that $u_i^* \in \partial_{F,\varepsilon_i} f(u_i)$ and $u_i \to_f x$ and $\varepsilon_i \downarrow 0$. Above, for $\varepsilon \geqslant 0$

$$\partial_{F,\varepsilon} f(u) = \left\{x^* \in X^*\colon\ \liminf_{h \to 0} \frac{f(u+h) - f(u) - \langle x^*, h\rangle}{\|h\|} \geqslant -\varepsilon\right\}$$

is the *Fréchet ε-subdifferential* of f at any u for which $f(u)$ is finite. We adopt the convention $\partial_{F,\varepsilon} f(u) = \emptyset$ when $|f(u)| = +\infty$. We also put $\partial_F f(u) = \partial_{F,0} f(u)$. If f is the indicator function ψ_S of a set $S \subset X$, that is, $\psi_S(u) = 0$ if $u \in S$ and $\psi_S(u) = +\infty$ otherwise, the set

$$N^{F,\varepsilon}(S, u) := \partial_{F,\varepsilon} \psi_S(u)$$

is called the ε-Fréchet normal set of S at u, and this set is obviously a cone when $\varepsilon = 0$. The *Mordukhovich limiting normal cone* $N^L(S,x)$ of S at x being defined by

$$N^L(S,x) := \partial_L \psi_S(x),$$

one has the equality

$$N^L(S,x) = w^*\text{-}\operatorname*{Lim\,sup}_{u \xrightarrow{S} x,\, \varepsilon \downarrow 0} N^{F,\varepsilon}(S,u).$$

If S is the epigraph of the function f, that is, $S = \operatorname{epi} f := \{(u,r) \in X \times \mathbb{R}\colon\ f(u) \leqslant r\}$, the subdifferentials $\partial_F f(x)$ and $\partial_L f(x)$ are related to the epigraph of f through the equalities

$$\begin{aligned} \partial_F f(x) &= \bigl\{x^* \in X^*\colon\ \bigl(x^*, -1\bigr) \in N^F(S,\zeta)\bigr\},\\ \partial_L f(x) &= \bigl\{x^* \in X^*\colon\ \bigl(x^*, -1\bigr) \in N^L(S,\zeta)\bigr\}, \end{aligned} \tag{2.3}$$

for $x \in X$ with $|f(x)| < +\infty$, and $\zeta := (x, f(x))$.

In the setting of *Asplund* space, that is, when X is a Banach space such that the topological dual of any separable subspace of X is separable, Mordukhovich and Shao [42] established the following characterization of the Mordukhovich limiting subdifferential

$$\partial_L f(x) = w^*\text{-}\operatorname*{Lim\,sup}_{u \xrightarrow{f} x} \partial_F f(u) \tag{2.4}$$

whenever the function f is lower semicontinuous near x.

It is known, see for example [41, Theorem 1.97] that for a closed set S of the Banach space X and $x \in S$ one has

$$N^L(S,x) = \mathbb{R}_+ \partial_L d_S(x), \tag{2.5}$$

where d_S denotes the distance function to the set S. It is also worth pointing out that

$$\partial_L d_S(x) = w^*\text{-}\operatorname*{Lim\,sup}_{u \xrightarrow{S} x,\, \varepsilon \downarrow 0} \partial_{F,\varepsilon} d_S(u) \tag{2.6}$$

whenever the space X is Asplund. This means that one may for $f = \psi_S$ in (2.2) require u to belong to the closed set S as $u \to x$.

When the space X is Asplund, f is finite at x and lower semicontinuous on an open neighborhood U of x and g is finite and Lipschitz continuous on U, then for any $\varepsilon > 0$ and $x' \in U$ and for any $x^* \in \partial_{F,\varepsilon}(f+g)(x')$ (see, for example, [41, Theorem 2.3]) there are $u, v \in B(x', \varepsilon) \cap U$ with $|f(u) - f(x')| < \varepsilon$ such that

$$x^* \in \partial_F f(u) + \partial_F g(v) + 2\varepsilon \mathbb{B}_{X^*}, \tag{2.7}$$

so that

$$\partial_L (f+g)(x) \subset \partial_L f(x) + \partial_L g(x). \tag{2.8}$$

The theory of Fréchet and limiting subdifferentials is developed in Mordukhovich's book [41].

One of the best ways to introduce the Clarke subdifferential (called also the Clarke generalized gradient) is to define it first for locally Lipschitz continuous functions via the generalized directional derivative. Recall that, for a locally Lipschitz continuous function $f : U \to \mathbb{R}$ on an open set U of a normed vector space X, its *Clarke generalized directional derivative* (see [13]) is defined for $x \in U$ by

$$f^o(x;h) := \limsup_{u \to x;\, t \downarrow 0} t^{-1}\big[f(u+th) - f(u)\big]$$

and then its *Clarke subdifferential* at x can be defined similarly to the convex setting by

$$\partial_C f(x) := \big\{x^* \in X^*\colon \langle x^*, h\rangle \leqslant f^o(x;h),\ \forall h \in X\big\}.$$

It is not difficult to see that for a locally Lipschitz continuous function f we have on one hand

$$f^o(x;-h) = (-f)^o(x;h), \tag{2.9}$$

and on the other hand

$$f^o(x;h) = \limsup_{(u,w) \to (x,h);\, t \downarrow 0} t^{-1}\big[f(u+tw) - f(u)\big],$$

and the latter equality entails that the function $f^o(\cdot;\cdot)$ is upper semicontinuous on $U \times X$, we refer to [13] for details. We also note that for any $h \in X$ and any dense set Q of X the set $\{(t,u) \in]0,+\infty[\times X\colon u + th \in Q\}$ is dense in $]0,+\infty[\times X$, so the continuity of f ensures

$$f^o(x;h) = \limsup_{\substack{u \to x;\, t \downarrow 0 \\ u+th \in Q}} t^{-1}\big[f(u+th) - f(u)\big]. \tag{2.10}$$

It is worth pointing out that, when X is Asplund, the Clarke subdifferential $\partial_C f(x)$ is related to the Mordukhovich limiting subdifferential $\partial_L f(x)$ through the equality

$$\partial_C f(x) = \overline{\mathrm{co}}^*\big(\partial_L f(x)\big), \tag{2.11}$$

where $\overline{\mathrm{co}}^*$ denotes the w^*-closed convex hull, we refer to [41, Theorem 3.57] for details.

The Clarke subdifferential obeys to the following sum rule (see, for example, [13, Corollary 1, p. 105]):

$$\partial_C(f+g)(x) \subset \partial_C f(x) + \partial_C g(x) \tag{2.12}$$

provided that one of the two functions $f, g : X \to \mathbb{R} \cup \{+\infty\}$ is locally Lipschitz continuous around x.

We know that the Clarke subdifferential of a locally Lipschitz continuous function on a normed space X is a nonempty weak* compact set of X^* and that the limiting subdifferential of a locally Lipschitz continuous function is also a nonempty (but it can be neither convex nor weak* closed) set provided that the space X is Asplund.

We recall that the function $f: X \to \mathbb{R}$ is *Gâteaux* (resp. *Hadamard*) *differentiable* at the point x of the normed space X provided there exists a continuous linear functional $A := D_G f(x) \in X^*$ (resp. $A =: D_H f(x)$) such that

$$\lim_{t\downarrow 0} t^{-1}\big[f(x+th) - f(x)\big] = \langle A, h\rangle \tag{2.13}$$

(resp. and provided the convergence is uniform for h in compact sets). If the convergence in (2.13) is uniform on every bounded subset, then we say that f is *Fréchet differentiable* and the derivative is denoted by $D_F f(x)$ in place of A. When in place of (2.13), one has

$$\lim_{\substack{u\to x\\ t\downarrow 0}} t^{-1}\big[f(u+th) - f(u)\big] = \langle A, h\rangle \tag{2.14}$$

with uniform convergence for h in compact sets of X, one says that G is *strictly Hadamard differentiable* at x and one puts $A =: D^s_H G(x)$. If the convergence in (2.14) is uniform for h in bounded sets of X, then f is said to be *strictly Fréchet differentiable* at x and one sets $A =: D^s_F G(x)$.

The lower Dini directional derivative of f at $u \in \operatorname{dom} f$ is given by

$$d^- f(u; h) := \liminf_{w\to h;\, t\downarrow 0} t^{-1}\big[f(u+tw) - f(u)\big]$$

and when f is Lipschitz continuous near u we obviously have for all $h \in X$

$$d^- f(u; h) = \liminf_{t\downarrow 0} t^{-1}\big[f(u+th) - f(u)\big].$$

The Dini subdifferential of f at x is the set

$$\partial^- f(x) = \big\{x^* \in X^*\colon \langle x^*, h\rangle \leqslant d^- f(x; h),\ \forall h \in X\big\}$$

for $x \in \operatorname{dom} f$ and $\partial^- f(x) = \emptyset$ if $x \notin \operatorname{dom} f$, where $\operatorname{dom} f := \{x \in X\colon |f(x)| < +\infty\}$ denotes the effective domain of f.

When f is Lipschitz continuous near u and

$$f^o(u; \cdot) = d^- f(u; \cdot)$$

the function f is said to be *Clarke directionally subregular at the point* u. This is easily seen to be equivalent to the existence of the (usual) directional derivative $f'(u; \cdot)$ and to the equality

$$f^o(u; \cdot) = f'(u; \cdot),$$

where

$$f'(u; h) := \lim_{t\downarrow 0} t^{-1}\big[f(u+th) - f(u)\big].$$

It is easy to see, for a closed set S of the normed space X and $x \in X$, that the directional derivatives of d_S and d_S^2 are linked as

$$\big(d_S^2\big)^\circ(x; h) = 2d_S(x)d_S^\circ(x; h), \qquad d^- d_S^2(x; h) = 2d_S(x)d^- d_S(x; h), \quad \forall h \in X. \tag{2.15}$$

As we see above, whenever nonconvex functions are considered, there are several ways to define their subdifferential. It is then natural to work with an abstract general concept of subdifferential allowing us to state many results in a general unified framework. Such a generality is achieved through the concept of presubdifferential. Following [53,54] a *presubdifferential* on a normed vector space X is an operator ∂ which assigns to any function $f : X \to \mathbb{R} \cup \{+\infty\}$ and any $x \in X$ a subset $\partial f(x)$ of X^* and which satisfies the following properties:

(P1) $\partial f(x) \subset X^*$ and $\partial f(x) = \emptyset$ if $x \notin \operatorname{dom} f$;
(P2) $\partial f(x) = \partial g(x)$ whenever f and g coincide on a neighborhood of x;
(P3) $\partial f(x)$ is equal to the subdifferential in the sense of Convex Analysis whenever f is convex and lower semicontinuous;
(P4) if f is lower semicontinuous near $x \in \operatorname{dom} f$, g is (finite) convex continuous near x, and x is a local minimum point of $f + g$, then one has

$$0 \in w^*\text{-}\operatorname*{Lim\,sup}_{u \xrightarrow{f} x} \partial f(u) + \partial g(x),$$

where $u \xrightarrow{f} x$ means $(u, f(u)) \to (x, f(x))$ and $w^*\text{-}\operatorname{Lim\,sup}_{u \xrightarrow{f} x} \partial f(u)$ denotes the weak* sequential outer (upper) limit of $\partial f(u)$ as defined in (2.2).

The graph of the presubdifferential set-valued mapping ∂f is the set

$$\operatorname{gph} \partial f := \bigl\{ \bigl(x, x^*\bigr) \in X \times X^* \colon x^* \in \partial f(x) \bigr\}.$$

When $f : U \to \mathbb{R} \cup \{+\infty\}$ is defined on a subset U of X its presubdifferential $\partial f(x)$ is defined as the presubdifferential of the extension of f to X with the value $+\infty$ outside of the set U. We mention that other abstract subdifferentials have been used in [30] in the study of sufficient conditions for the metric regularity of set-valued mappings.

We say that ∂ is a subdifferential with exact inclusion sum rule when (P1)–(P3) hold and instead of (P4) one requires:

(P4′) for any function g finite and locally Lipschitz continuous near x

$$\partial(f + g)(x) \subset \partial f(x) + \partial g(x)$$

and $0 \in \partial f(x)$ whenever $x \in \operatorname{dom} f$ is a local minimum of f.

The above mentioned (Clarke, Fréchet, Mordukhovich) subdifferentials are presubdifferentials in appropriate spaces. It is known that the presubdifferential of a function does not determine the function up to a constant. In order to have this integration property we have to distinguish a proper class of functions. The largest class of functions was given in [54], see also [34]. Before recalling the definition, we denote, by $\varphi'(t; 1)$ the right derivative of a function φ at t defined on an interval of $\mathbb{R}$, whenever it exists, that is,

$$\varphi'(t; 1) := \lim_{\tau \downarrow 0} \frac{\varphi(t + \tau) - \varphi(t)}{\tau}.$$

Definition 2.1. (See [54].) Let U be a nonempty open convex subset of a Banach space X and $g : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function on U with $U \cap \operatorname{dom} g \neq \emptyset$ and $\mu > 0$ be fixed. Let D be a subset of X with $\operatorname{Dom} \partial g \subset D \subset \operatorname{dom} g$. We say that the function g is essentially ∂, μ-directionally smooth (∂, μ-eds, for short) on U relative to D provided that for each $u \in U \cap \operatorname{Dom} \partial g$:

(i) for each $v \in U \cap \operatorname{dom} g$ the function $g_{u,v}(t) := g(u + t(v - u))$ is finite and continuous on $[0, 1]$;
(ii) for each $v \in U \cap D$ there are real numbers $0 = t_0 < \cdots < t_p = 1$ such that the function $t \mapsto g_{u,v}(t)$ is absolutely continuous on each closed interval included in $[0, 1] \setminus \{t_0, t_1, \ldots, t_p\}$;
(iii) for each $v \in U \cap D$ with $v \neq u$ there exists a subset $T \subset [0, 1]$ of full Lebesgue measure (that is, of Lebesgue measure 1) such that for every $t \in T$ and every sequence $\{(x_i, x_i^*)\}_{i\in\mathbb{N}}$ in $\operatorname{gph} \partial g$ with $x_i \to x(t) := u + t(v - u)$, there is some $w \in]x(t), v]$ for which

$$\limsup_{i\to\infty} \langle x_i^*, w - x_i\rangle \leqslant \|w - x(t)\| \left(\|v - u\|^{-1} g'_{u,v}(t; 1) + \mu\right).$$

When the above properties hold for all $\mu > 0$, one says that g is ∂-eds on U relative to D. If $D = \operatorname{dom} g$, one merely says that g is ∂-eds on U.

Let U be a nonempty open convex subset of a Banach space X and let $f : U \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function which is not identically equal to $+\infty$. Then each one of the following conditions ensures (see [54]) that f is ∂-eds on U with respect to $D = \operatorname{dom} f$:

(a) f is convex on U;
(b) f is locally Lipschitz continuous on U and segment-wise essentially smooth on U, and the presubdifferential ∂f is included in the Clarke subdifferential of f (we recall that if for each $(u, v) \in U \times U$ and for $x(t) = u + t(v - u)$ the set

$$\left\{t \in]0, 1[: \ f^o\left(x(t); -x'(t)\right) \neq -f^o\left(x(t); x'(t)\right)\right\}$$

has null Lebesgue measure, then the function f is said to be *segment-wise essentially smooth on* U);
(c) f is locally Lipschitz continuous on U and directionally subregular on U, and the presubdifferential ∂f is included in the Clarke subdifferential of f;
(d) ∂ is a subdifferential included in the Clarke subdifferential with exact subdifferential sum rule (P4$'$) and f is locally DC on U, that is, for any $u \in U$ there exist an open convex neighborhood $U' \subset U$ of u, a lower semicontinuous convex function $f_1 : U' \to \mathbb{R} \cup \{+\infty\}$ and a continuous convex function $f_2 : U' \to \mathbb{R}$ such that $f(x) = f_1(x) - f_2(x)$ for all $x \in U'$.

Further, f is ∂-eds on U relative to $D := \operatorname{Dom} \partial f$ whenever ∂ is included in the Clarke subdifferential and f is approximate convex on U in the sense that for every $u \in U$ and every $\varepsilon > 0$ there exists some convex neighborhood $U' \subset U$ of u such that for all $x, x' \in U'$ and all $t \in]0, 1[$ one has

$$f\left(tx + (1 - t)x'\right) \leqslant tf(x) + (1 - t)f\left(x'\right) + \varepsilon t(1 - t)\left\|x' - x\right\|.$$

3. Subregularity properties of the Moreau envelope

When we deal with functions which are not sufficiently smooth then we face with several difficulties, for example: exploring monotonicity properties by means of "derivative" needs advanced knowledge from nonsmooth analysis, identification of functions (up to an additive constant) by its "derivative" needs distinguishing a proper class of functions for which it is possible, we refer to [54] for recent achievements in that direction. However, approximation techniques allow us to look for a close function with some better properties, for example some convex functions can be identified by means of the Moreau envelope, see for example [59]. Below, under the assumption that X is a Banach space with uniformly Gâteaux differentiable norm (this property is enjoyed by: Hilbert space, separable space – it has an equivalent norm with this property, L_p with $1 < p < \infty$, super-reflexive space), see [21–23,31,32,52], we prove that the Moreau envelope of a function is a ∂-eds function on X with respect to the whole space. In other words, under the additional assumptions that the function is lower semicontinuous and bounded from below by a negative quadratic function, see Definition 3.1 below, this result guarantees that there is a function from the class of functions with integrable subdifferential, which is close to f in some sense, since for such a function f the equality

$$\operatorname{epi} f = \bigcap_{\lambda > 0} \operatorname{epi} e_\lambda f$$

holds true, keep also in mind that the inclusion $\operatorname{epi} f \subset \operatorname{epi} e_\lambda f$ holds true for every $\lambda > 0$ and the reverse inclusion "almost" holds true on bounded sets, see (3.8) below for details.

Let X be a normed vector space and $Q \subset X$ be a dense subset. For any function $f : X \to \mathbb{R} \cup \{+\infty\}$, any $\lambda > 0$ and any $\varepsilon \geqslant 0$, let us put

$$\begin{aligned} e_\lambda f(x) &:= \inf_{z \in X} f(z) + \frac{1}{2\lambda}\|z - x\|^2, \\ P_\lambda f(x) &:= \arg\min_{z \in X}\left\{ f(z) + \frac{1}{2\lambda}\|z - x\|^2 \right\}, \\ \mathcal{E}_{\lambda,\varepsilon,Q} f(x) &:= \left\{ z \in X\colon\ \exists y \in Q \text{ s.t. } e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z - y\|^2 \text{ and } \|x - y\| \leqslant \varepsilon \right\} \end{aligned} \tag{3.1}$$

and

$$A(f, x, \lambda, \varepsilon, Q) := \Big\{ y^* \in X^*\colon\ \exists y \in Q,\ \exists z \in X \text{ s.t. } y^* \in \partial\big((2\lambda)^{-1}\|\cdot\|^2\big)(y - z), \\ e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z - y\|^2 \text{ and } \|x - y\| \leqslant \varepsilon \Big\}$$

and

$$E_{\lambda,Q} f(x) := \bigcap_{\epsilon > 0} \operatorname{cl}_{w^*} A(f, x, \lambda, \varepsilon, Q),$$

where cl_{w^*} stands for the weak* closure the closure of the set with respect to the weak* topology.

The set $A(f,x,\lambda,\varepsilon,Q)$ above involves the subdifferential of the convex function $\frac{1}{2}\|\cdot\|^2$. It is known (and easily seen) that for all x in the normed vector space X

$$\partial\left(\frac{1}{2}\|\cdot\|^2\right)(x)=\left\{x^*\in X^*\colon \langle x^*,x\rangle=\|x\|^2,\ \|x^*\|=\|x\|\right\}, \tag{3.2}$$

and the graph of this set-valued mapping is $\|\cdot\|\times w(X^*,X)$ closed in $X\times X^*$. Further, the set $\partial(\frac{1}{2}\|\cdot\|^2)(x)$ is a singleton for all $x\in X$ if and only if the norm $\|\cdot\|$ is Gâteaux differentiable outside zero, and in such a case writing $\partial(\frac{1}{2}\|\cdot\|^2)(x)=\{J(x)\}$ the mapping

$$J:X\to X^* \text{ is norm-to-weak}^* \text{ continuous}, \tag{3.3}$$

that is, the mapping J (usually called the duality (single-valued) mapping) is continuous with respect to the norm topology on X and the weak* topology on X^*. Obviously, one has

$$J(tx)=tJ(x)\quad \text{for all } t\in\mathbb{R},\ x\in X. \tag{3.4}$$

When the Moreau envelope is Gâteaux differentiable, its Gâteaux derivative is easily seen to be connected with the duality mapping as it is shown in Proposition 3.1. Of course, a similar statement also holds true for the Dini subdifferential (and others) when $e_\lambda f$ is not Gâteaux differentiable, see, for example, [12,14]. Before stating the result, recall that an extended real-valued function f on X is *proper* whenever f is finite at some point of X and does not take on the value $-\infty$ at any point of X.

Proposition 3.1. *Let $(X,\|\cdot\|)$ be a normed space whose norm $\|\cdot\|$ is Gâteaux differentiable outside zero and $f:X\to\mathbb{R}\cup\{+\infty\}$ be a proper function. Let a real $\lambda>0$ be given. Assume that $x\in X$ is such that the Gâteaux derivative $D_Ge_\lambda f(x)$ exists and $P_\lambda f(x)\neq\emptyset$. Then*

$$\forall z\in P_\lambda f(x),\quad D_Ge_\lambda f(x)=\frac{1}{\lambda}J(x-z)\quad \textit{and}\quad \langle D_Ge_\lambda f(x),x-z\rangle=\frac{1}{\lambda}\|x-z\|^2.$$

Proof. Let $x\in X$ be such that the Gâteaux derivative $D_Ge_\lambda f(x)$ exists and $P_\lambda f(x)\neq\emptyset$. We have

$$\begin{aligned}\langle D_Ge_\lambda f(x),h\rangle&=\lim_{t\downarrow 0}\frac{e_\lambda f(x)-e_\lambda f(x-th)}{t}\\&\geqslant \lim_{t\downarrow 0}\frac{f(z)+\frac{1}{2\lambda}\|x-z\|^2-f(z)-\frac{1}{2\lambda}\|x-th-z\|^2}{t}=\left\langle\frac{1}{\lambda}J(x-z),h\right\rangle\end{aligned}$$

for every $h\in X$, $z\in P_\lambda f(x)$, which combined with (3.2) implies the statement. □

It is easy to observe that for any $\lambda>0$, $\varepsilon>0$ and $x\in X$, the sets $A(f,x,\lambda,\varepsilon,Q)$ are nonempty whenever $e_\lambda f(x)$ is finite. It is also interesting to know whether there are $x\in X$ such that the functions $f(\cdot)+\frac{1}{2\lambda}\|\cdot-x\|^2$ attain their infimum, that is, when $P_\lambda f(x)\neq\emptyset$. Of course in the finite-dimensional setting $P_\lambda f(x)\neq\emptyset$ for every $x\in X$ and λ sufficiently small, whenever f is lower semicontinuous and bounded from below by some negative quadratic function. On the contrary, when X is a general infinite-dimensional normed space, we do not expect that

$P_\lambda f(x) \neq \emptyset$ for every $x \in X$. Moreover, it may even happen for some function f that $P_\lambda f(x) = \emptyset$ for every $x \in X$, whenever X is not reflexive. This is a consequence of the proposition below, where it is shown that the space X is reflexive whenever for any $x^* \in X^*$ taken in place of f there exists some $x \in X$ such that the infimum in the definition of $e_1 x^*(x)$ is attained. This shows the need for some functions f not to consider, in the definition of sets $A(f, x, \lambda, \varepsilon, Q)$, only those y for which infima of $f(\cdot) + \frac{1}{2\lambda}\|\cdot - y\|^2$ are attained.

Proposition 3.2. *Assume that $(X, \|\cdot\|)$ is a Banach space such that for every $x^* \in X^*$ there is $x \in X$ such that for some $z \in X$ we have*

$$e_1 x^*(x) = \langle x^*, z\rangle + \frac{1}{2}\|z - x\|^2.$$

Then X is a reflexive Banach space.

Proof. In order to prove the reflexivity it is enough to show that the closed unit ball $\mathbb{B}_X$ is weakly compact, see [29, Section 16] for example. For this reason we use the theorem of James, see [29, Section 19]. Let us fix $x^* \in X^* \setminus \{0\}$, $x \in X$ and $z \in X$ such that

$$\inf_{u \in X} \langle x^*, u\rangle + \frac{1}{2}\|u - x\|^2 = \langle x^*, z\rangle + \frac{1}{2}\|z - x\|^2.$$

By a simple subdifferential calculus we obtain $-x^* \in \partial(\frac{1}{2}\|\cdot - x\|^2)(z)$, which yields by (3.2)

$$\langle -x^*, z - x\rangle = \|z - x\|^2 \quad \text{and} \quad \|x^*\| = \|z - x\|.$$

Hence

$$\langle x^*, \|x - z\|^{-1}(x - z)\rangle = \|z - x\|,$$

so x^* attains its maximum on $\mathbb{B}_X$ at $\|x - z\|^{-1}(x - z)$. Thus every continuous linear functional attains its maximum on the closed ball $\mathbb{B}_X$, so the ball is weakly compact and the space is reflexive. $\square$

The following definition is needed for guaranteeing the local Lipschitzness of $e_\lambda f$ as well as the boundedness of $E_{\lambda,Q} f(x)$.

Definition 3.1. We say that f is *bounded from below by a negative quadratic function* if there exist $\gamma \in \mathbb{R}$, $\beta \geqslant 0$, and $\alpha \geqslant 0$ such that

$$f(x) \geqslant -\alpha\|x\|^2 - \beta\|x\| + \gamma, \quad \forall x \in X. \tag{3.5}$$

We may easily see that this definition is equivalent to say that there exists $\alpha_0 \geqslant 0$ such that

$$f(x) \geqslant -\alpha_0\big(\|x\|^2 + 1\big), \quad \forall x \in X. \tag{3.6}$$

Consider a proper convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ bounded from below by a quadratic function $q(\cdot)$. Then the lower semicontinuous hull $\bar{f}$ of f is bounded from below by $q(\cdot)$ and hence it is a proper lower semicontinuous convex function. Consequently $\bar{f}$ is bounded from below by a continuous affine function. So

$$\begin{aligned}&f \text{ is bounded from below by a quadratic function}\\ &\quad\text{if and only if it is bounded from below by a continuous affine function.}\end{aligned} \tag{3.7}$$

We establish now the boundedness of $E_{\lambda,Q}f(x)$ and the local Lipschitz property of $e_\lambda f$. An earlier result concerning the Lipschitz property of $e_\lambda f$ over a ball $r\mathbb{B}_X$ can be found in [3, Theorem 2.64] with different constants. An Attouch–Wets convergence property is also proved.

Proposition 3.3. *Let $(X, \|\cdot\|)$ be a normed vector space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function bounded from below by a negative quadratic function, and $Q \subset X$ be a given dense subset. Then the following hold:*

(a) *The sets $\mathcal{E}_{\lambda,\varepsilon,Q}f(x)$ are nonempty and bounded, and the sets $\mathrm{cl}_{w^*} A(f, x, \lambda, \varepsilon, Q)$ and $E_{\lambda,Q}f(x)$ are nonempty and weak* compact for all $x \in X$ and $\lambda \in]0, \frac{1}{2\alpha}[$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$).*
(b) *For each real $\lambda \in]0, \frac{1}{2\alpha}[$ the function $e_\lambda f$ is Lipschitz continuous on each ball $r\mathbb{B}_X$ of X with a Lipschitz constant therein $L \geqslant \frac{r}{\lambda}$.*
(c) *For each $c > 0$ and $\varepsilon > 0$ there is $\eta > 0$ such that for all $\lambda \in]0, \eta[$ the following inclusion*

$$\operatorname{epi} e_\lambda f \cap B[0, c] \subset \operatorname{epi} f + B(0, \varepsilon) \tag{3.8}$$

holds true; and this combined with the inclusion $\operatorname{epi} f \subset \operatorname{epi} e_\lambda f$ (due to the inequality $e_\lambda f \leqslant f$) ensures the Attouch–Wets (see [51]) convergence of $\{e_\lambda f\}_\lambda$ to f as $\lambda \downarrow 0$.

Proof. Let us fix $\lambda \in]0, \frac{1}{2\alpha}[$ and $x \in X$. First observe that $e_\lambda f(x)$ is finite, due to the relation (3.5) and the properness of f. Now fix $\varepsilon > 0$ and take $y \in B[x, \varepsilon]$. Observe that if

$$e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z - y\|^2$$

then, by (3.5), we get

$$e_\lambda f(x) + \varepsilon \geqslant -\alpha\|z\|^2 - \beta\|z\| + \gamma + \frac{1}{2\lambda}\|z - y\|^2.$$

So

$$e_\lambda f(x) + \varepsilon - \gamma \geqslant -\alpha\|z\|^2 - \beta\|z\| + \frac{1}{2\lambda}\|z\|^2 + \frac{1}{2\lambda}\|y\|^2 - \frac{1}{\lambda}\|z\|\|y\|. \tag{3.9}$$

Thus

$$e_\lambda f(x) + \varepsilon - \gamma \geqslant \left(\frac{1}{2\lambda} - \alpha\right)\|z\|^2 - \left(\beta + \frac{\|x\| + \varepsilon}{\lambda}\right)\|z\|,$$

which implies the existence of $M_1 > 0$, not depending on y, such that if $y \in B[x, \varepsilon]$ and $e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z - y\|^2$ then

$$z \in B(0, M_1),$$

in other words

$$\mathcal{E}_{\lambda,\varepsilon,Q} f(x) \subset \mathcal{E}_{\lambda,\varepsilon,X} f(x) \subset B(0, M_1).$$

Finally, there exists $M_2 > 0$ such that if $y^* \in \partial((2\lambda)^{-1}\|\cdot\|^2)(y - z)$ and $e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z - y\|^2$, with $\|x - y\| \leqslant \varepsilon$, then $y^* \in B(0, M_2)$. So for all $\varepsilon' \in]0, \varepsilon]$ the sets $A(f, x, \lambda, \varepsilon', Q)$ are bounded since they are included in $B(0, M_2)$. The nonemptiness of all of them follows from the fact that for $y = x$ there exists $z' \in X$ such that

$$e_\lambda f(x) + \varepsilon > f(z') + \frac{1}{2\lambda}\|z' - y\|^2,$$

so by the density of Q we get

$$e_\lambda f(x) + \varepsilon > f(z') + \frac{1}{2\lambda}\|z' - q\|^2$$

for some $q \in Q$. Hence the sets $\mathrm{cl}_{w^*} A(f, x, \lambda, \varepsilon, Q)$ are weak* compact and thus the intersection $\bigcap_{0<\varepsilon'\leqslant\varepsilon} \mathrm{cl}_{w^*} A(f, x, \lambda, \varepsilon, Q)$ is nonempty and weak* compact, which asserts that $E_{\lambda,Q} f(x)$ is nonempty and weak* compact. The assertion (a) is established.

Concerning (b) fix a point $z_0 \in X$ where f is finite and a real $r > 0$. For $\varepsilon > 0$, $x \in r\mathbb{B}_X$, and $z \in X$ such that $e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|x - z\|^2$, writing

$$f(z_0) + \frac{1}{\lambda}\left(r^2 + \|z_0\|^2\right) \geqslant f(z_0) + \frac{1}{2\lambda}\|x - z_0\|^2 \geqslant e_\lambda f(x)$$

and choosing $y = x$ in (3.9) yields

$$f(z_0) + \frac{1}{\lambda}\left(r^2 + \|z_0\|^2\right) + \varepsilon - \gamma \geqslant \left(\frac{1}{2\lambda} - \alpha\right)\|z\|^2 - \left(\beta + \frac{r}{\lambda}\right)\|z\|,$$

which gives some real $M > 0$ depending only on λ and r such that $\|z\| \leqslant M$. This ensures that $e_\lambda f(x) = \inf_{z\in M\mathbb{B}_X}[f(z) + \frac{1}{2\lambda}\|x - z\|^2]$ for all $x \in r\mathbb{B}_X$. Observing for each $z \in M\mathbb{B}_X$ that the function $\frac{1}{2\lambda}\|\cdot - z\|^2$ is Lipschitz continuous on $r\mathbb{B}_X$ with $\frac{1}{\lambda}(r + M)$ as Lipschitz constant therein the latter inequality for $e_\lambda f(x)$ guarantees $|e_\lambda f(x) - e_\lambda f(x')| \leqslant \frac{1}{\lambda}(r + M)\|x - x'\|$ for all $x, x' \in r\mathbb{B}_X$.

In order to prove (c) we can endow $X \times \mathbb{R}$ with the norm $\|(u, r)\| := \max\{\|u\|, |r|\}$. Let $\alpha_0 > 0$ be as in relation (3.6) and let $c > 0$ and $\varepsilon > 0$ be arbitrary. Let $\lambda > 0$ and $(x', r) \in \operatorname{epi} e_\lambda f \cap B[0, c]$. Then there exists $u \in X$ such that

$$f(u) + \frac{1}{2\lambda}\|u - x'\|^2 \leqslant e_\lambda f(x') + \lambda \leqslant r + \lambda \leqslant c + \lambda. \tag{3.10}$$

As

$$\|u\|^2 \leqslant 2\|u - x'\|^2 + 2\|x'\|^2 \leqslant 2\|u - x'\|^2 + 2c^2$$

we obtain, by using relations (3.6) and (3.10),

$$-2\|u-x'\|^2 \leqslant -\|u\|^2 + 2c^2 \leqslant 1 + \frac{1}{\alpha_0} f(u) + 2c^2$$
$$\leqslant 1 + \frac{1}{\alpha_0}\left[c + \lambda - \frac{1}{2\lambda}\|u-x'\|^2\right] + 2c^2.$$

Thus $\|u-x'\|^2(\frac{1}{2\lambda\alpha_0} - 2) \leqslant 1 + 2c^2 + \frac{c+\lambda}{\alpha_0}$, which, for $0 < \lambda < 1/(4\alpha_0)$, is equivalent to

$$\|u-x'\|^2 \leqslant (\alpha_0 + 2\alpha_0 c^2 + c + \lambda)\frac{2\lambda}{1-4\alpha_0\lambda}.$$

Choose a positive real $\eta < \min(\frac{1}{4\alpha_0}, \varepsilon)$ such that $(\alpha_0 + 2\alpha_0 c^2 + c + \lambda)\frac{2\lambda}{1-4\alpha_0\lambda} < \varepsilon^2$ for all $0 < \lambda < \eta$. Then, for any fixed number $\lambda \in]0, \eta[$, we have $\|u - x'\|^2 \leqslant \varepsilon^2$. So, $(x', r) \in (u, r + \varepsilon) + B[0, \varepsilon]$, with $(u, r + \varepsilon) \in \operatorname{epi} f$ (by relation (3.10)). Consequently

$$\operatorname{epi} e_\lambda f \cap B[0, c] \subset \operatorname{epi} f + B[0, \varepsilon], \quad \forall \lambda \in]0, \eta[. \qquad \square$$

Remark 3.1. Let X be a normed vector space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. For all $x \in \operatorname{dom} f$ there is $\delta_x > 0$ such that the function $(f + \psi_{B[x,\delta_x]})(\cdot)$ is bounded from below (if the function f is bounded from below then we put $\delta_x := \infty$, $B[x, \infty] := X$), where we recall that

$$\psi_S(y) := \begin{cases} 0 & \text{if } y \in S, \\ +\infty & \text{if } y \notin S. \end{cases}$$

Hence the function $(f + \psi_{B[x,\delta_x]})(\cdot)$ satisfies the assumptions of the above proposition. In particular for all nonempty subsets $S \subset X$ we have

$$e_\lambda \psi_S(x) = \frac{d_S^2(x)}{2\lambda} = \frac{\inf_{s\in S}\|x-s\|^2}{2\lambda}.$$

Of course if f is the sum of a Lipschitz continuous function on X and a proper convex and lower semicontinuous function, then the assumptions of the above proposition are also satisfied.

Let us recall that a norm $\|\cdot\|$ on a vector space X is said to be uniformly Gâteaux differentiable (off zero) in a direction $h \in X$ if for any $\varepsilon > 0$, there exists a real $\delta(h, \varepsilon) > 0$ such that for every $u \in X$ with $\|u\| = 1$, there is a continuous linear functional f_u on X for which

$$\left|\frac{\|u+th\| - \|u\|}{t} - f_u(h)\right| < \varepsilon$$

for all $t \in]0, \delta(h, \varepsilon)[$ or equivalently (see [17, Definition 1.1, Definition 6.5 and Lemma 6.6])

$$\lim_{t\downarrow 0} \sup_{u\in X,\ \|u\|=1} t^{-1}(\|u-th\| + \|u+th\| - 2) = 0 \quad \text{for all } h \in X.$$

Of course, in such a case $f_u(h) = \langle J(u), h\rangle$.

Remark 3.2. Usually when we want to apply the above equality we are not in the unit sphere. Moreover it is often more convenient to use the norm squared instead of the norm itself. Under the assumption that the norm $\|\cdot\|$ on the vector space X is uniformly Gâteaux differentiable (off zero) in a direction $h \in X$, for all reals $0 < \rho < r$ we have

$$\lim_{t \downarrow 0} \sup_{y \in X,\ \rho \leqslant \|y\| \leqslant r} t^{-1}\bigl(\|y + th\|^2 - \|y\|^2 - 2\langle J(y), th\rangle\bigr) = 0.$$

Indeed, for $0 < t < 1$, putting $u := \|y\|^{-1}y$ and $s := s(t, y) := \|y\|^{-1}t$ (depending on both t and $\|y\|$) we have

$$\begin{aligned}
& t^{-1}\bigl(\|y + th\|^2 - \|y\|^2 - 2t\langle J(y), h\rangle\bigr) \\
&\quad = \|y\|\bigl\{s^{-1}\bigl(\|u + sh\|^2 - \|u\|^2 - 2s\langle J(u), h\rangle\bigr)\bigr\} \\
&\quad = \|y\|\bigl\{s^{-1}\bigl(\|u + sh\| - \|u\| - s\langle J(u), h\rangle\bigr)\bigl(\|u + sh\| + \|u\|\bigr) \\
&\qquad + \bigl(\|u + sh\| + \|u\| - 2\bigr)\langle J(u), h\rangle\bigr\} \\
&\quad = \|y\|\bigl\{s^{-1}\bigl(\|u + sh\| - \|u\| - s\langle J(u), h\rangle\bigr)\bigl(\|u + sh\| + \|u\|\bigr) \\
&\qquad + \bigl(\|u + sh\| - \|u\|\bigr)\langle J(u), h\rangle\bigr\} \\
&\quad \leqslant \|y\|\bigl(2 + s\|h\|\bigr)s^{-1}\bigl(\|u + sh\| - \|u\| - s\langle J(u), h\rangle\bigr) + \|y\|s\|h\|\bigl|\langle J(u), h\rangle\bigr| \\
&\quad \leqslant \bigl(2r + \|h\|\bigr)s^{-1}\bigl(\|u + sh\| - \|u\| - s\langle J(u), h\rangle\bigr) + t\|h\|\bigl|\langle J(u), h\rangle\bigr|.
\end{aligned}$$

Since

$$\sup_{v \in X,\ \|v\|=1} \tau^{-1}\bigl(\|v + \tau h\| - \|v\| - \tau\langle J(v), h\rangle\bigr) \to 0 \quad \text{as } \tau \downarrow 0$$

and since $\sup_{y \in X,\ \rho \leqslant \|y\| \leqslant r} s(t, y) \to 0$ as $t \downarrow 0$, we deduce that

$$\lim_{t \downarrow 0} \sup_{y \in X,\ \rho \leqslant \|y\| \leqslant r} t^{-1}\bigl(\|y + th\|^2 - \|y\|^2 - 2t\langle J(y), h\rangle\bigr) = 0$$

(because we know that $t^{-1}(\|y + th\|^2 - \|y\|^2 - 2t\langle J(y), h\rangle) \geqslant 0$ for all $t > 0$). This justifies the desired equality.

More generally, the above arguments show the following: For any family $(Y_t)_{0<t<1}$ of subsets of a fixed ball $r\mathbb{B}_X$ such that $\inf_{y \in Y_t}(\|y\|^{-1}t) \to 0$ as $t \downarrow 0$, the property

$$\lim_{t \downarrow 0} \sup_{y \in Y_t} t^{-1}\bigl(\|y + th\|^2 - \|y\|^2 - 2t\langle J(y), h\rangle\bigr) = 0$$

is satisfied.

The norm is uniformly Gâteaux differentiable when it is uniformly Gâteaux differentiable in each direction $h \in X$. An equivalent characterization for the norm of a Banach space to be uniformly Gâteaux differentiable in a direction h, by means of the directional derivative of the distance function, was given by L. Zajiček, see [63, Theorem 3].

Theorem 3.1. *Let $(X, \|\cdot\|)$ be a Banach space and $h \in X$. Then the following conditions are equivalent*:

(i) *The norm $\|\cdot\|$ of X is uniformly Gâteaux differentiable in the direction $h \in X$.*
(ii) *For any nonempty closed subset $S \subset X$ and any $x \notin S$ the directional derivative $\lim_{t\downarrow 0} \frac{d_S(x+th)-d_S(x)}{t}$ exists.*

It is shown in the theorem below that whenever the norm of the vector space X is uniformly Gâteaux differentiable, then the Moreau envelope enjoys remarkable properties as regards its directional derivatives. In the theorem a general formula for the Clarke directional derivative is also established in the case when the proximal mapping $P_\lambda f(x)$ may be empty, see (3.11) below. As it will be observed in Remark 3.3 the formula yields the lower semicontinuity of the lower Dini directional derivative of the Moreau envelope, that is, the lower semicontinuity of $d^-(e_\lambda f)(\cdot;\cdot)$. It is worth pointing out that our method for obtaining the properties of directional derivatives of the Moreau envelope does not require any semicontinuity assumption on f.

Theorem 3.2. *Let $(X, \|\cdot\|)$ be a normed vector space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function* (*i.e.* $\operatorname{dom} f \neq \emptyset$). *Suppose that f is bounded from below by a negative quadratic function, where α is as in relation (3.5)* (*with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$*). *Let a positive $\lambda \in]0, \frac{1}{2\alpha}[$ be given. Then for every dense subset $Q \subset X$ the following hold*:

The function $e_\lambda f$ is locally Lipschitz continuous and for all $x, h \in X$

$$d^-(-e_\lambda f)(x; -h) = \lim_{t\uparrow 0} \frac{e_\lambda f(x+th) - e_\lambda f(x)}{t} = (e_\lambda f)^o(x; h)$$
$$= \max_{u^*\in E_{\lambda,X} f(x)} \langle u^*, h\rangle = \max_{u^*\in E_{\lambda,Q} f(x)} \langle u^*, h\rangle, \tag{3.11}$$

$$\lim_{t\downarrow 0} \frac{e_\lambda f(x+th) - e_\lambda f(x)}{t} = \min_{u^*\in E_{\lambda,X} f(x)} \langle u^*, h\rangle = \min_{u^*\in E_{\lambda,Q} f(x)} \langle u^*, h\rangle \tag{3.12}$$

and

$$\partial_C(e_\lambda f)(x) = \overline{\mathrm{co}}^*\big(E_{\lambda,X} f(x)\big) = \overline{\mathrm{co}}^*\big(E_{\lambda,Q} f(x)\big) = -\partial^-(-e_\lambda f)(x), \tag{3.13}$$

where $\overline{\mathrm{co}}^*(\cdot)$ denotes the weak* closed convex hull.

Proof. The assertion (b) of Proposition 3.3 ensures that $e_\lambda f(\cdot)$ is locally Lipschitz continuous on X. Fix $h, x \in X$ and a dense subset $Q \subset X$. By Proposition 3.3 the set $E_{\lambda,Q} f(x)$ is nonempty and weak* compact in X^*. First, we remark that

$$\liminf_{t\uparrow 0} t^{-1}\big[e_\lambda f(x+th) - e_\lambda f(x)\big] = \liminf_{t\downarrow 0} t^{-1}\big[e_\lambda f(x) - e_\lambda f(x-th)\big]$$
$$= d^-(-e_\lambda f)(x; -h)$$

and the following inequalities hold true

$$\begin{aligned}\liminf_{t\downarrow 0} t^{-1}\big[e_\lambda f(x) - e_\lambda f(x-th)\big] &\leqslant \limsup_{t\downarrow 0} t^{-1}\big[e_\lambda f(x) - e_\lambda f(x-th)\big] \\ &\leqslant \limsup_{t\downarrow 0} t^{-1}\big[e_\lambda f\big(z(t)+th\big) - e_\lambda f\big(z(t)\big)\big] \\ &\leqslant (e_\lambda f)^o(x;h) \end{aligned} \tag{3.14}$$

where $z(t) = x - th$. To get (3.11) it remains to prove that

$$\liminf_{t\downarrow 0} t^{-1}\big[e_\lambda f(x) - e_\lambda f(x-th)\big] \geqslant \max_{u^*\in E_{\lambda,Q} f(x)} \langle u, h\rangle \geqslant (e_\lambda f)^o(x;h).$$

Indeed, if the last one is true then

$$\liminf_{t\uparrow 0} t^{-1}\big[e_\lambda f(x+th) - e_\lambda f(x)\big] \geqslant (e_\lambda f)^o(x;h) \geqslant \limsup_{t\uparrow 0} t^{-1}\big[e_\lambda f(x+th) - e_\lambda f(x)\big]$$

and this ensures that $\lim_{t\uparrow 0} t^{-1}[e_\lambda f(x+th) - e_\lambda f(x)]$ exists.

Fix any real $r > \liminf_{t\downarrow 0} t^{-1}[e_\lambda f(x) - e_\lambda f(x-th)]$. There exist $t_i \in]0,1]$ such that $t_i \to 0$ and

$$r > t_i^{-1}\big[e_\lambda f(x) - e_\lambda f(x - t_i h)\big].$$

Let $u^* \in E_{\lambda,X} f(x)$. Then for every $\varepsilon > 0$

$$\begin{aligned}\langle u^*, h\rangle \in \mathrm{cl}_{\mathbb{R}}\Big\{\langle y^*, h\rangle\colon\ &\exists y, z\in X \text{ s.t. } y^* \in \partial\big((2\lambda)^{-1}\|\cdot\|^2\big)(y-z), \\ &e_\lambda f(x) + \varepsilon \geqslant f(z) + \frac{1}{2\lambda}\|z-y\|^2 \text{ and } \|x-y\| \leqslant \varepsilon\Big\}.\end{aligned}$$

So, there exist sequences $\{y_i\}_{i\in\mathbb{N}}$, $\{z_i\}_{i\in\mathbb{N}}$ in X and $\{u_i^*\}_{i\in\mathbb{N}}$ in X^* such that $\langle u_i^*, h\rangle \to \langle u^*, h\rangle$ and such that for all $i \in \mathbb{N}$

$$\begin{aligned} &e_\lambda f(x) + t_i^2 \geqslant f(z_i) + \frac{1}{2\lambda}\|z_i - y_i\|^2, \\ &\|y_i - x\| \leqslant t_i^2 \quad \text{and} \quad u_i^* \in \partial\big((2\lambda)^{-1}\|\cdot\|^2\big)(y_i - z_i). \end{aligned} \tag{3.15}$$

Since $t_i \in]0,1]$ we have $\{z_i\}_{i\in\mathbb{N}} \subset \mathcal{E}_{\lambda,1,X} f(x)$ (see (3.1)) and hence, by Proposition 3.3, there exists $M > 1$ (not depending on t_i) such that $\max\{\|y_i - z_i\|^2, \|y_i - z_i - t_i h\|^2\} \leqslant M$ for all i. So for all i

$$\begin{aligned} r &> t_i^{-1}\bigg[-t_i^2 + f(z_i) + \frac{1}{2\lambda}\|z_i - y_i\|^2 - f(z_i) - \frac{1}{2\lambda}\|z_i - x + t_i h\|^2\bigg] \\ &= t_i^{-1}\bigg[-t_i^2 + \frac{1}{2\lambda}\|z_i - y_i\|^2 - \frac{1}{2\lambda}\|z_i - x + t_i h\|^2\bigg] \\ &= -t_i + t_i^{-1}\bigg[\frac{1}{2\lambda}\|z_i - y_i\|^2 - \frac{1}{2\lambda}\big\|(z_i - y_i + t_i h) + y_i - x\big\|^2\bigg] \end{aligned}$$

$$\geqslant -t_i + t_i^{-1}\left[\frac{1}{2\lambda}\left(\|y_i - z_i\|^2 - \|y_i - z_i - t_i h\|^2\right)\right] - t_i \frac{2M}{\lambda}$$
$$= -t_i + \langle u_i^*, h\rangle + t_i^{-1}\left[\frac{1}{2\lambda}\left(2\|y_i - z_i\|^2 - \|y_i - z_i - t_i h\|^2 - \|y_i - z_i + t_i h\|^2\right)\right]$$
$$+ t_i^{-1}\left[\frac{1}{2\lambda}\left(\|y_i - z_i + t_i h\|^2 - \|y_i - z_i\|^2\right) - \langle u_i^*, t_i h\rangle\right] - t_i \frac{2M}{\lambda}.$$

Since $\frac{1}{2\lambda}(\|y_i - z_i + t_i h\|^2 - \|y_i - z_i\|^2) - \langle u_i^*, t_i h\rangle \geqslant 0$ because $u_i^* \in \partial((2\lambda)^{-1}\|\cdot\|^2)(y_i - z_i)$, we deduce

$$r \geqslant -t_i + \langle u_i^*, h\rangle + t_i^{-1}\left[\frac{1}{2\lambda}\left(2\|y_i - z_i\|^2 - \|y_i - z_i - t_i h\|^2 - \|y_i - z_i + t_i h\|^2\right)\right]$$
$$- t_i \frac{2M}{\lambda}. \tag{3.16}$$

Observing that

$$t_i^{-1}\left(2\|y_i - z_i\|^2 - \|y_i - z_i - t_i h\|^2 - \|y_i - z_i + t_i h\|^2\right) \geqslant -4\|h\|\|y_i - z_i\| - 2t_i\|h\|^2,$$

we see that, if there is a subsequence of the sequence $\{y_i - z_i\}_{i\in\mathbb{N}}$ converging to 0, then we get

$$r \geqslant \langle u^*, h\rangle.$$

Let us consider the case $\liminf_{i\to\infty}\|y_i - z_i\| > 0$. Put $u_i := \|y_i - z_i\|^{-1}(y_i - z_i)$ and $s_i := \|y_i - z_i\|^{-1}t_i$ for all $i \in \mathbb{N}$. Note that $s_i \to 0$ and write

$$A_i := t_i^{-1}\left(2\|y_i - z_i\|^2 - \|y_i - z_i - t_i h\|^2 - \|y_i - z_i + t_i h\|^2\right)$$
$$= \left(\|y_i - z_i\|\right)s_i^{-1}\left(2 - \|u_i - s_i h\|^2 - \|u_i + s_i h\|^2\right).$$

Remark 3.2 ensures that $s_i^{-1}(2 - \|u_i - s_i h\|^2 - \|u_i + s_i h\|^2) \to 0$, hence $A_i \to 0$ according to the boundedness of $\{\|y_i - z_i\|\}_i$. It then follows from (3.16) that

$$r \geqslant \langle u^*, h\rangle.$$

This establishes the inequality

$$\liminf_{t\downarrow 0} t^{-1}\left[e_\lambda f(x) - e_\lambda f(x - th)\right] \geqslant \max_{u^*\in E_{\lambda,X} f(x)}\langle u^*, h\rangle. \tag{3.17}$$

Now, we will show that

$$(e_\lambda f)^o(x; h) \leqslant \max_{u\in E_{\lambda,Q} f(x)}\langle u^*, h\rangle. \tag{3.18}$$

So let us choose by (2.10) sequences $\{t_i\}_{i\in\mathbb{N}} \subset]0,\infty[$, $\{x_i\}_{i\in\mathbb{N}} \subset X$ such that $t_i \to 0$, $x_i + t_i h \in Q$ for every $i \in \mathbb{N}$, $x_i \to x$ and

$$(e_\lambda f)^o(x; h) = \lim_{i\to\infty} t_i^{-1}\left[e_\lambda f(x_i + t_i h) - e_\lambda f(x_i)\right].$$

For every $i \in \mathbb{N}$ take according to the definition of $e_\lambda f(x_i)$ some $z_i \in X$ such that

$$e_\lambda f(x_i) + t_i^2 \geqslant f(z_i) + \frac{1}{2\lambda}\|x_i - z_i\|^2. \tag{3.19}$$

Denoting by γ a Lipschitz constant of $e_\lambda f$ on some neighborhood U of x, we may suppose (without loss of generality) that $x_i \in U$ for all integers i, and hence we have

$$e_\lambda f(x_i) + t_i^2 \leqslant e_\lambda f(x) + t_i^2 + \gamma\|x_i - x\|.$$

Putting

$$\varepsilon_i := t_i^2 + \max\{1, \gamma\}\|x_i - x\| + \frac{t_i}{\lambda}\|h\|\|x_i - z_i\| + \frac{1}{2\lambda}t_i^2\|h\|^2 + t_i\|h\|$$

and using (3.19) we see that

$$\begin{aligned} f(z_i) + \frac{1}{2\lambda}\|x_i + t_i h - z_i\|^2 &\leqslant f(z_i) + \frac{1}{2\lambda}\|x_i - z_i\|^2 + \frac{t_i}{\lambda}\|h\|\|x_i - z_i\| + \frac{1}{2\lambda}t_i^2\|h\|^2 \\ &\leqslant e_\lambda f(x) + \varepsilon_i. \end{aligned}$$

Further, on one hand $\varepsilon_i \to 0$ since $\{x_i - z_i\}_{i\in\mathbb{N}}$ is bounded according to (3.19) and to the minorization of f by a quadratic function, and on the other hand

$$\|x_i + t_i h - x\| \leqslant \|x_i - x\| + t_i\|h\| \leqslant \varepsilon_i.$$

Consequently, for $u_i^* = D_G((2\lambda)^{-1}\|\cdot\|^2)(x_i - z_i + t_i h)$, we obtain that $u_i^* \in A(f, x, \lambda, \varepsilon_i, Q)$. Denoting (see Proposition 3.3) by u_0^* a weak* cluster point of $\{u_i^*\}_{i\in\mathbb{N}}$, we have $u_0^* \in E_{\lambda,Q} f(x)$ (because $\varepsilon_i \to 0$), and by (3.19) and Remark 3.2

$$\begin{aligned} (e_\lambda f)^o(x; h) &= \lim_{i\to\infty} \frac{1}{t_i}\big[e_\lambda f(x_i + t_i h) - e_\lambda f(x_i)\big] \\ &\leqslant \liminf_{i\to\infty} \frac{1}{t_i}\left[t_i^2 + \frac{1}{2\lambda}\|x_i - z_i + t_i h\|^2 - \frac{1}{2\lambda}\|x_i - z_i\|^2\right] \\ &\leqslant \liminf_{i\to\infty}\langle u_i^*, h\rangle \leqslant \langle u_0^*, h\rangle \leqslant \max_{u^*\in E_{\lambda,Q} f(x)}\langle u^*, h\rangle, \end{aligned}$$

that is, (3.18) holds true.

The equalities $\max_{u^*\in E_{\lambda,Q} f(x)}\langle u^*, h\rangle = \max_{u^*\in E_{\lambda,X} f(x)}\langle u^*, h\rangle$ and

$$(e_\lambda f)^o(x; h) = \lim_{t\uparrow 0}\frac{e_\lambda f(x + th) - e_\lambda f(x)}{t} = \max_{u^*\in E_{\lambda,X} f(x)}\langle u^*, h\rangle$$

follow from relations (3.14), (3.17) and (3.18). Changing h into $-h$ yields

$$\lim_{t\uparrow 0}\frac{e_\lambda f(x - th) - e_\lambda f(x)}{t} = \max_{u^*\in E_{\lambda,Q} f(x)}\langle u^*, -h\rangle = -\min_{u^*\in E_{\lambda,Q} f(x)}\langle u^*, h\rangle,$$

which translates (3.12).

Finally, observing that

$$\begin{aligned}\max\{\langle u^*, h\rangle\colon u^* \in \overline{\mathrm{co}}^*\big(E_{\lambda,Q} f(x)\big)\} &= \max\{\langle u^*, h\rangle\colon u^* \in E_{\lambda,Q} f(x)\} \\ &= \max\{\langle u^*, h\rangle\colon u^* \in E_{\lambda,X} f(x)\} \\ &= \max\{\langle u^*, h\rangle\colon u^* \in \overline{\mathrm{co}}^*\big(E_{\lambda,X} f(x)\big)\},\end{aligned}$$

and keeping in mind (3.11) we deduce the equality

$$-\partial^-(-e_\lambda f)(x) = \partial_C(e_\lambda f)(x) = \overline{\mathrm{co}}^*\big(E_{\lambda,X} f(x)\big) = \overline{\mathrm{co}}^*\big(E_{\lambda,Q} f(x)\big),$$

which completes the proof. □

It has been observed by R.T. Rockafellar and R.J.-B. Wets (see book [51, Example 10.32]) that the opposite of the Moreau envelope is a lower-C^2 function, whenever the space $(X, \|\ \|)$ is a finite-dimensional Euclidean space. Their arguments use the local compactness of the space and the Rademacher theorem to express (3.11) and (3.12) in terms of $P_\lambda f(x)$, which is a nonempty set in this case; we refer to [51, Theorem 9.61, Theorem 10.31, Example 10.32] for details. The nonemptiness of the sets $P_\lambda f(x)$ for x in some dense subset of X, has been also used recently by R. Cibulka and M. Fabian (see [12]) to obtain a description of the Clarke subdifferential of $e_\lambda f$ for a function f fulfilling the property (3.20) below in a Banach space with uniformly Gâteaux differentiable norm. In order to discuss their result and compare it with Theorem 3.2 above let us recall their statement.

Theorem 3.3. *Let X be a Banach space whose norm is uniformly Gâteaux differentiable, let $f : X \to \mathbb{R} \cup \{+\infty\}$ be either Lipschitzian on X, or be proper, bounded from below, and lower semicontinuous and such that the set*

$$A_f := \big\{y \in X \;\big|\; \arg\min f(\cdot) + \|\cdot - y\|^2 \neq \emptyset\big\} \tag{3.20}$$

is dense in X, and let $x \in X$. Consider any subset $M \subset A_f$ that is dense in a neighborhood of x and denote by $D_{f,x,M}$ the set of possible weak limits of all sequences $\partial(\|\cdot\|^2)(x_i - z_i)$, where $\{x_i\}_{i\in\mathbb{N}} \subset M$ is converging to x and $\{z_i\}_{i\in\mathbb{N}}$ is such that the minimum of the function $f(\cdot) + \|\cdot - x_i\|^2$ is attained at z_i for each $i \in \mathbb{N}$.*

(i) *If $e_{\frac{1}{2}} f(x) < f(x)$, then $\partial_C e_{\frac{1}{2}} f(x) = \overline{\mathrm{co}}^* D_{f,x,M}$.*
(ii) *If $e_{\frac{1}{2}} f(x) = f(x)$, then $\partial_C e_{\frac{1}{2}} f(x) = \overline{\mathrm{co}}^*(\{0\} \cup D_{f,x,M})$.*

It follows from Proposition 3.2 that for every nonreflexive Banach space with uniformly Gâteaux differentiable norm there exists $x^* \in X^*$ such that the set A_{x^*} is empty. Because of that, if the Banach space X is not reflexive, there are even smooth functions for which Theorem 3.3 cannot be applied, while Theorem 3.2 can be applied to any Lipschitz continuous function or any proper, bounded from below one, whenever the norm $\|\cdot\|$ of the normed vector space is uniformly Gâteaux differentiable. However, if all the assumptions of Theorem 3.3 are satisfied we infer by (3.11), (3.12) and (3.13) that for any dense set Q of X we have:

(i′) If $e_{\frac{1}{2}} f(x) < f(x)$, then $\partial_C e_{\frac{1}{2}} f(x) = \overline{\text{co}}^* D_{f,x,M} = \overline{\text{co}}^* E_{\frac{1}{2},Q} f(x)$.
(ii′) If $e_{\frac{1}{2}} f(x) = f(x)$, then $\partial_C e_{\frac{1}{2}} f(x) = \overline{\text{co}}^*(\{0\} \cup D_{f,x,M}) = \overline{\text{co}}^* E_{\frac{1}{2},Q} f(x)$.

Using Zajiček's result in Theorem 3.1, we know that, in the case of Banach space, Theorem 3.2 cannot be extended to spaces where the norm is not uniformly Gâteaux differentiable. This is stated in the theorem below.

Theorem 3.4. *For a normed vector space* $(X, \|\cdot\|)$ *consider the following properties*:

(a) *the norm* $\|\cdot\|$ *of* X *is uniformly Gâteaux differentiable*;
(b) *for each proper lower semicontinuous function* $f : X \to \mathbb{R} \cup \{+\infty\}$, *which is bounded from below by a negative quadratic function, the opposite of its Moreau envelope* $e_\lambda f$ *is Clarke directionally subregular for all* $\lambda \in]0, \frac{1}{2\alpha}[$, *where* α *is as in relation* (3.5).

Then the implication (a) $\Rightarrow$ (b) *holds true, and this implication is an equivalence whenever the space* $(X, \|\cdot\|)$ *is a Banach space.*

Proof. (a) $\Rightarrow$ (b) Theorem 3.2 asserts that for $x, h \in X$

$$\lim_{t \uparrow 0} \frac{e_\lambda f(x - th) - e_\lambda f(x)}{t} = (e_\lambda f)^o(x; -h).$$

Now invoking the known fact that $(-e_\lambda f)^o(x; h) = (e_\lambda f)^o(x; -h)$, see (2.9), we get

$$\lim_{t \downarrow 0} \frac{(-e_\lambda f)(x + th) - (-e_\lambda f)(x)}{t} = (e_\lambda f)^o(x; -h) = (-e_\lambda f)^o(x; h).$$

(b) $\Rightarrow$ (a) It follows from Theorem 3.1 that if for every closed subset $S \subset X$, its distance function $x \mapsto d_S(x)$ (associated with the norm $\|\cdot\|$) always has right-hand Gâteaux directional derivatives at every point off the set, that is, for each $x \notin S$,

$$d'_S(x, h) := \lim_{t \downarrow 0} \frac{d_S(x + th) - d_S(x)}{t}$$

exists for all $h \in X$, then the norm $\|\cdot\|$ is uniformly Gâteaux differentiable.

To establish our implication, it suffices to show that, for each closed nonempty set S and for each $x \notin S$ the function $h \mapsto d'_S(x, h)$ is well defined. Indeed, let f be the indicator function of S, that is, $f(x) = 0$ if $x \in S$ and $f(x) = \infty$ if $x \notin S$, that is, $f = \psi_S$. Then for all $\lambda \in]0, \infty[$ we have

$$e_\lambda f(u) = \frac{1}{2\lambda} d_S^2(u), \quad \forall u \in X.$$

By (b) the function $-e_\lambda f$ is Clarke directionally subregular, that is,

$$(e_\lambda f)'(x; h) := \lim_{t \downarrow 0} \frac{e_\lambda f(x + th) - e_\lambda f(x)}{t} = -(-e_\lambda f)^o(x; h), \quad \forall h \in X$$

or equivalently

$$\left(d_S^2\right)'(x;h) = -\left(-d_S^2\right)^o(x;h), \quad \forall h \in X. \tag{3.21}$$

Now invoking the equalities in (2.15), we get that

$$\begin{aligned} \left(d_S^2\right)'(x;h) &= 2d_S(x)d_S'(x;h), \\ \left(-d_S^2\right)^o(x;h) = \left(d_S^2\right)^o(x;-h) &= 2d_S(x)d_S^o(x;-h), \quad \forall h \in X. \end{aligned}$$

Combining the last equalities with that in (3.21) and taking into account the equality $d^o(x;-h) = (-d_S^2)^o(x;h)$, we get

$$(-d_S)'(x;h) = (-d_S)^o(x;h)$$

so the right-hand Gâteaux directional derivative $d_S'(x;h)$ exists. □

Remark 3.3. Let us observe that under the assumptions of Theorem 3.2, for every $\lambda \in]0, \frac{1}{2\alpha}[$ (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$), where α is as in relation (3.5), Theorem 3.4 above ensures that the function $-e_\lambda f(\cdot)$ is directionally subregular, so it is ∂-eds on X for any presubdifferential included in the Clarke subdifferential (see the example (c) following Definition 2.1). In particular for every subset $S \subset X$ the function $-d_S^2(\cdot)$ is directionally subregular and ∂-eds on X. In fact, for every $x \in X$, $h \in X$ we have

$$(-e_\lambda f)^o(x;h) = (e_\lambda f)^o(x;-h) = d^-(-e_\lambda f)(x;h), \tag{3.22}$$

which gives the directional subregularity and consequently the function $-e_\lambda f(\cdot)$ is ∂-eds on X. The directional subregularity of $-d_S^2(\cdot)$ is a consequence of the equality $e_{\frac{1}{2}}\psi_S(x) = d_S^2(x)$, see Remark 3.1. Moreover it follows from (3.22) and the upper semicontinuity of $(e_\lambda f)^o(\cdot;\cdot)$ that the function $d^-(e_\lambda f)(\cdot;\cdot)$ is lower semicontinuous on the product $X \times X$. Indeed the equality between the first and third member in (3.22) guarantees that the usual directional derivative of $-e_\lambda f$ at x exists and hence the equalities

$$d^-(e_\lambda f)(x;h) = -d^-(-e_\lambda f)(x;h) = -(-e_\lambda f)^o(x;h)$$

hold true. So, the upper semicontinuity of $(-e_\lambda f)^o(\cdot;\cdot)$ allows us to conclude.

Now we can use Theorem 3.2 to obtain through the relation (3.11) the following result:

Theorem 3.5. *Suppose that all the assumptions of Theorem 3.2 are satisfied and let $x \in X$. The following properties are equivalent*:

(a) $\partial^- e_\lambda f(x) \neq \emptyset$;
(b) $\partial_C e_\lambda f(x)$ *is a singleton*;
(c) $e_\lambda f$ *is strictly Hadamard differentiable at* x;
(d) $\partial_C e_\lambda f(x) = \partial^- e_\lambda f(x) = \{D_H e_\lambda f(x)\}$;
(e) $\partial_C e_\lambda f(x) = \partial^- e_\lambda f(x)$.

Suppose in addition that the function f is weakly sequentially lower semicontinuous at each point of its effective domain, the space $(X, \|\cdot\|)$ is Hilbert and $P_\lambda f(x) \neq \emptyset$. Then each of the above conditions is equivalent to the following property:

(f) *$e_\lambda f$ is strictly Fréchet differentiable at x.*

Proof. Let $x^* \in \partial^- e_\lambda f(x) \neq \emptyset$. Then, by Theorem 3.2, we have

$$\liminf_{t\downarrow 0} \frac{e_\lambda f(x+th) - e_\lambda f(x)}{t} = \min_{u^*\in\partial_C e_\lambda f(x)} \langle u^*, h\rangle, \quad \forall h \in X$$

and hence for all $u^* \in \partial_C e_\lambda f(x)$

$$\langle x^*, h\rangle \leqslant \liminf_{t\downarrow 0} \frac{e_\lambda f(x+th) - e_\lambda f(x)}{t} \leqslant \langle u^*, h\rangle, \quad \forall h \in X.$$

Thus for all $u^* \in \partial_C e_\lambda f(x)$

$$\langle x^*, h\rangle \leqslant \langle u^*, h\rangle, \quad \forall h \in X$$

or equivalently $u^* = x^*$, and hence $\partial_C e_\lambda f(x) = \{x^*\}$. So the implication (a) $\Rightarrow$ (b) is established.

The implication (b) $\Rightarrow$ (c) follows from Proposition 2.2.4 in [13] which says that a locally Lipschitz continuous function g is strictly Hadamard differentiable at x if and only if $\partial_C g(x)$ is a singleton. The implications (c) $\Rightarrow$ (d) and (d) $\Rightarrow$ (e) are obvious, and the implication (e) $\Rightarrow$ (a) is due to the nonemptiness of $\partial_C e_\lambda f(x)$ according to the Lipschitz property of $e_\lambda f$ near x. The implication (f) $\Rightarrow$ (c) is obvious.

It remains to prove the converse implication (c) $\Rightarrow$ (f) under the additional assumptions of the theorem. So, suppose that the space X is a Hilbert space and $P_\lambda f(x) \neq \emptyset$. First observe that it follows from (3.11) and from item (d) that

$$E_{\lambda,X} f(x) = \{D_G e_\lambda f(x)\}. \tag{3.23}$$

Consider any sequence $\{y_i\}_{i\in\mathbb{N}}$ converging to x and any sequence $\{z_i\}_{i\in\mathbb{N}}$ in X such that $e_\lambda f(x) = \lim_{i\to\infty} f(z_i) + \frac{1}{2\lambda}\|y_i - z_i\|^2$. The equality (3.23) and the definition of $E_{\lambda,X} f(x)$ yield $\frac{1}{\lambda} J(y_i - z_i) \xrightarrow{w} D_G e_\lambda f(x)$ (in the weak topology, which is the same as weak* topology since we are in the case of Hilbert space). Let $w \in X$ satisfying $J(x - w) = \lambda D_G e_\lambda f(x)$. Take any $z \in P_\lambda f(x)$. By Proposition 3.1 we have $D_G e_\lambda f(x) = \frac{1}{\lambda} J(x-z)$, so $\frac{1}{\lambda} J(x-w) = \frac{1}{\lambda} J(x-z)$ thus $w = z$ and $\frac{1}{\lambda} J(x - z_i) \xrightarrow{w} \frac{1}{\lambda} J(x - w)$. The weak lower semicontinuity of f and $\|x - \cdot\|^2$ at $z \in P_\lambda f(x) \subset \operatorname{dom} f$, and the equality $z = w$ give

$$e_\lambda f(x) = \lim_{i\to\infty} f(z_i) + \frac{1}{2\lambda}\|y_i - z_i\|^2 \geqslant f(w) + \frac{1}{2\lambda}\|x - w\|^2 \geqslant e_\lambda f(x),$$

hence $\lim_{i\to\infty} \|y_i - z_i\| = \|x - w\|$ (according to the weak lower semicontinuity again of f and $\|\cdot\|$), so $\lim_{i\to\infty} \|w - z_i\| = 0$. Consequently, $P_\lambda f(x) = \{w\}$ is a singleton and the whole sequence $\{z_i\}_{i\in\mathbb{N}}$ converges strongly to w.

Now take any sequences $\{x_i\}_{i\in\mathbb{N}}$ and $\{h_i\}_{i\in\mathbb{N}}$ in X such that $x_i \to x$ and $h_i \to 0$ with $h_i \neq 0$. Choose $\{a_i\}_{i\in\mathbb{N}}$ and $\{b_i\}_{i\in\mathbb{N}}$ in X such that

$$e_\lambda f(x_i) + \|h_i\|^2 \geqslant f(a_i) + \frac{1}{2\lambda}\|x_i - a_i\|^2 \geqslant e_\lambda f(x_i),$$
$$e_\lambda f(x_i + h_i) + \|h_i\|^2 \geqslant f(b_i) + \frac{1}{2\lambda}\|x_i + h_i - b_i\|^2 \geqslant e_\lambda f(x_i + h_i)$$

for every $i \in \mathbb{N}$. On one hand, the latter inequalities imply

$$e_\lambda f(x) = \lim_{i\to\infty} f(a_i) + \frac{1}{2\lambda}\|x_i - a_i\|^2 \quad \text{and} \quad e_\lambda f(x) = \lim_{i\to\infty} f(b_i) + \frac{1}{2\lambda}\|x_i + h_i - b_i\|^2,$$

which combined with what precedes gives

$$\lim_{i\to\infty}\|w - a_i\| = \lim_{i\to\infty}\|w - b_i\| = 0.$$

On the other hand, the same inequalities also guarantee that

$$\begin{aligned}
-\|h_i\| + \left\langle \frac{1}{\lambda\|h_i\|} J(x_i - b_i), h_i \right\rangle &\leqslant \frac{-\|h_i\|^2 + \frac{1}{2\lambda}\|x_i + h_i - b_i\|^2 - \frac{1}{2\lambda}\|x_i - b_i\|^2}{\|h_i\|} \\
&\leqslant \frac{e_\lambda f(x_i + h_i) - e_\lambda f(x_i)}{\|h_i\|} \\
&\leqslant \frac{\|h_i\|^2 + \frac{1}{2\lambda}\|x_i + h_i - a_i\|^2 - \frac{1}{2\lambda}\|x_i - a_i\|^2}{\|h_i\|} \\
&\leqslant \|h_i\| + \left\langle \frac{1}{\lambda\|h_i\|} J(x_i + h_i - a_i), h_i \right\rangle.
\end{aligned}$$

Further, $x_i + h_i - a_i \to x - w$ and $x_i - b_i \to x - w$, hence we have $\frac{1}{\lambda}J(x_i - b_i) \to \frac{1}{\lambda}J(x - w)$ and $\frac{1}{\lambda}J(x_i + h_i - a_i) \to \frac{1}{\lambda}J(x - w)$, and this along with the last inequalities implies

$$\lim_{\|h\|\downarrow 0,\, x'\to x} \frac{e_\lambda f(x' + h) - e_\lambda f(x') - \langle D_G e_\lambda f(x), h\rangle}{\|h\|} = 0,$$

which translates the property (f). □

Let us point out that a function can be weakly sequentially lower semicontinuous at each point of its effective domain but it may fail to be weakly sequentially lower semicontinuous on X. Indeed, taking any set S of an infinite-dimensional normed space X which is not weakly sequentially closed, the indicator function ψ_S is weakly lower semicontinuous on its effective domain S but it is not weakly sequentially lower semicontinuous on X.

It is known that there are Clarke directionally subregular functions which are not strictly Hadamard differentiable. However, if we impose suitable assumptions on the space as in Theorem 3.2, then we get through the equivalence between (a) and (c) of the theorem above that the Moreau envelope is Clarke directionally subregular if and only if it is strictly Hadamard differentiable.

Corollary 3.1. *Under the assumptions of Theorem 3.2, the function $e_\lambda f$ is Clarke directionally subregular at x if and only if it is strictly Hadamard differentiable at the point x.*

Recalling that continuous convex functions are Clarke directionally subregular at any point of its effective domain, so for every proper convex function f bounded from below by a quadratic function the function $e_\lambda f$ is Clarke directionally subregular for every $\lambda > 0$ since such a function f is bounded from below by a continuous affine function (see (3.7)). Hence as a direct consequence of Corollary 3.1 we have

Corollary 3.2. *Let $(X, \|\cdot\|)$ be a normed vector space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function bounded from below by a quadratic function. Then for every real $\lambda > 0$ the function $e_\lambda f$ is strictly Hadamard differentiable at every $x \in X$.*

In (3.3) we already observed that the Gâteaux differentiability of the norm ensures the norm-to-weak* continuity of the duality mapping J. In Theorem 3.6 below we have much more under the additional uniform Gâteaux differentiability of the norm. More precisely, if $(X, \|\cdot\|)$ is a Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable, then the mapping J is continuous on each sphere with respect to the weak sequential convergence induced on the sphere and the weak* topology of X^*. Before proving the result, recall that the norm $\|\cdot\|$ has the (*sequential*) *Kadec–Klee property provided that, given a sequence $(x_i)_i$ in X, the convergences $x_i \xrightarrow{w} x$ ($\xrightarrow{w}$ stands for the convergence in the weak topology) and* $\|x_i\| \to \|x\|$ imply that $\|x_i - x\| \to 0$ as $i \to \infty$, see [17, Definition 1.1(iii), p. 42] where this property is defined in an equivalent form. Any uniformly convex norm fulfills this property (of course, the norm associated with the inner product of a Hilbert space has the property too). Let us also observe that the associated duality mapping

$$J : X \to X^* \text{ is continuous with respect to the norms } \|\cdot\| \text{ on } X \text{ and } \|\cdot\|_* \text{ on } X^*, \tag{3.24}$$

whenever the space is reflexive with its norm Gâteaux differentiable off the origin and the dual norm $\|\cdot\|_*$ of X^* has the Kadec–Klee property, since the convergence $x_i \to x$ ensures $\|J(x_i)\|_* = \|x_i\| \to \|x\| = \|J(x)\|_*$ and combining this with (3.3), the Kadec–Klee property gives $\|J(x_i) - J(x)\|_* \to 0$. Moreover, the continuity of J with respect to the norms $\|\cdot\|$ on X and $\|\cdot\|_*$ on X^* guarantees the C^1 property of the square of the norm.

The other property of the mapping J established in Theorem 3.6 below will be useful for the proof of Theorem 3.8. In the proof of Theorem 3.6 we use a subdifferential approximation technique, which to the best of our knowledge was first introduced by H. Attouch, see [3]; for some comments and extensions we refer to [62].

Theorem 3.6. *Let $(X, \|\cdot\|)$ be a Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable. Then for any $x \in X$ and any sequence $\{x_i\}_{i\in\mathbb{N}}$ such that $\|x_i\| \to \|x\|$ and $\liminf_{i\to\infty}\langle J(x), x_i\rangle \geqslant \|x\|^2$ we have*

$$J(x_i) \xrightarrow{w^*} J(x).$$

Proof. For $x = 0$ the sequence $\{x_i\}_{i\in\mathbb{N}}$ converges to 0 in norm and hence in that case the result follows from the norm to weak* continuity of J (see (3.3)). Now fix $x \in X \setminus \{0\}$ and a sequence $\{x_i\}_{i\in\mathbb{N}}$ such that $\|x_i\| \to \|x\|$, along with $\liminf_{i\to\infty}\langle J(x), x_i\rangle \geqslant \|x\|^2$. Define

$$f_i(z) := \frac{1}{2}\|x_i + z\|^2 - \frac{1}{2}\|x_i\|^2 - \langle J(x), z\rangle + \frac{1}{2}\|z\|^2 \quad \text{for all } z \in X.$$

Obviously $f_i(0) = 0$. Further, for every $z \in X$, since

$$\frac{1}{2}\|x_i + z\|^2 - \langle J(x), x_i + z\rangle + \frac{1}{2}\|x\|^2 \geqslant \frac{1}{2}\big(\|x_i + z\| - \|x\|\big)^2 \geqslant 0,$$

we have

$$\frac{1}{2}\|x_i + z\|^2 - \langle J(x), z\rangle \geqslant -\frac{1}{2}\|x\|^2 + \langle J(x), x_i\rangle.$$

Hence

$$f_i(z) \geqslant \frac{1}{2}\|z\|^2 - \frac{1}{2}\|x\|^2 - \frac{1}{2}\|x_i\|^2 + \langle J(x), x_i\rangle.$$

Take a sequence $\{\varepsilon_i\}_{i\in\mathbb{N}} \subset]0, 1[$ with $\varepsilon_i \downarrow 0$ and notice that by a form of the Ekeland variational principle we infer the existence of a sequence $\{z_i\}_{i\in\mathbb{N}} \subset X$ for which for each $i \in \mathbb{N}$ we have $f_i(z_i) \leqslant f_i(0) = 0$ and

$$f_i(z_i) \leqslant f_i(z) + \varepsilon_i\|z - z_i\| \quad \text{for every } z \in X. \tag{3.25}$$

Hence

$$0 \geqslant f_i(z_i) \geqslant \frac{1}{2}\|z_i\|^2 - \frac{1}{2}\|x\|^2 - \frac{1}{2}\|x_i\|^2 + \langle J(x), x_i\rangle,$$

which together with the inequality $\liminf_{i\to\infty}\langle J(x), x_i\rangle - \frac{1}{2}\|x\|^2 - \frac{1}{2}\|x_i\|^2 \geqslant 0$ implies $\frac{1}{2}\|z_i\|^2 \to 0$ as $i \to \infty$. Taking into account the Gâtaeux differentiability of the square of the norm we see that f_i is Gâteaux differentiable and $D_G f_i(z_i) = \{J(x_i + z_i) - J(x) + J(z_i)\}$ for all $i \in \mathbb{N}$. The convergence $\frac{1}{2}\|z_i\|^2 \to 0$ as $i \to \infty$ forces $\|J(z_i)\| \to 0$, as $i \to \infty$, and since z_i are minimizers of the convex functions $f_i(\cdot) + \varepsilon_i\|\cdot - z_i\|$ (see (3.25)), applying the "convex" subdifferential calculus we arrive at

$$\big\|J(x_i + z_i) - J(x)\big\| \to 0. \tag{3.26}$$

Take any subsequence $\{z_{i_k}\}_{k\in\mathbb{N}}$ such that $\|z_{i_k}\| > 0$ for every $k \in \mathbb{N}$ (if there is no such a subsequence then $\|J(x_i) - J(x)\| \to 0$) and put $t_k := \sqrt{\|z_{i_k}\|}$. Fix any $h \in X$. From the uniform Gâteaux differentiability of the norm $\|\cdot\|$ and from Remark 3.2 we have $\varepsilon_k \to 0$, where

$$\varepsilon_k := \max\left\{\left|\frac{\frac{1}{2}\|x_{i_k} + z_{i_k} + t_k h\|^2 - \frac{1}{2}\|x_{i_k} + z_{i_k}\|^2 - \langle J(x_{i_k} + z_{i_k}), t_k h\rangle}{t_k}\right|, \right.$$
$$\left.\left|\frac{\frac{1}{2}\|x_{i_k} + t_k h\|^2 - \frac{1}{2}\|x_{i_k}\|^2 - \langle J(x_{i_k}), t_k h\rangle}{t_k}\right|\right\}.$$

We also have

$$\begin{aligned}
t_k\big\langle J(x_{i_k}), h\big\rangle &\leqslant \frac{1}{2}\|x_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}\|^2\\
&\leqslant \frac{1}{2}\|x_{i_k}+z_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}+z_{i_k}\|^2\\
&\quad+\frac{1}{2}\|x_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}+z_{i_k}+t_kh\|^2+\frac{1}{2}\|x_{i_k}+z_{i_k}\|^2-\frac{1}{2}\|x_{i_k}\|^2\\
&\leqslant t_k\varepsilon_k+t_k\big\langle J(x_{i_k}+z_{i_k}), h\big\rangle\\
&\quad+\frac{1}{2}\|z_{i_k}\|\big(\|x_{i_k}+t_kh\|+\|x_{i_k}+z_{i_k}+t_kh\|+\|x_{i_k}+z_{i_k}\|+\|x_{i_k}\|\big)
\end{aligned}$$

and similarly

$$\begin{aligned}
t_k\big\langle J(x_{i_k}+z_{i_k}), h\big\rangle &\leqslant \frac{1}{2}\|x_{i_k}+z_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}+z_{i_k}\|^2\\
&\leqslant \frac{1}{2}\|x_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}\|^2\\
&\quad-\frac{1}{2}\|x_{i_k}+t_kh\|^2+\frac{1}{2}\|x_{i_k}+z_{i_k}+t_kh\|^2-\frac{1}{2}\|x_{i_k}+z_{i_k}\|^2+\frac{1}{2}\|x_{i_k}\|^2\\
&\leqslant t_k\varepsilon_k+t_k\big\langle J(x_{i_k}), h\big\rangle\\
&\quad+\frac{1}{2}\|z_{i_k}\|\big(\|x_{i_k}+t_kh\|+\|x_{i_k}+z_{i_k}+t_kh\|+\|x_{i_k}+z_{i_k}\|+\|x_{i_k}\|\big).
\end{aligned}$$

Thus $|\langle J(x_{i_k}), h\rangle-\langle J(x_{i_k}+z_{i_k}), h\rangle|\to 0$, which, by (3.26), implies $|\langle J(x_{i_k}), h\rangle-\langle J(x), h\rangle|\to 0$. It follows $J(x_i)\xrightarrow{w^*} J(x)$ and the proof is completed. □

The following result shows that Moreau envelope is ∂-eds for any presubdifferential ∂ included in the Clarke subdifferential.

Theorem 3.7. *Under the assumptions of Theorem 3.2, the Moreau envelope $e_\lambda f$ is ∂-eds for any presubdifferential ∂ included in the Clarke subdifferential and $D=X$.*

Proof. The local Lipschitz property of $e_\lambda f$ ensures on the one hand that conditions (i) and (ii) of Definition 2.1 are fulfilled, and on the other hand that for $u, v\in X$ with $v\neq u$ the function $t\mapsto e_\lambda f(u+t(v-u))$ is derivable on a subset $T\subset[0,1]$ of full measure in $[0,1]$. Take any $t\in T$ and any sequence $\{(x_k, x_k^*)\}_{k\in\mathbb{N}}$ in $\operatorname{gph}\partial e_\lambda f$ such that $x_k\to x(t):=u+t(v-u)$, and put $w:=v$, $h:=v-u$. It follows from the definition of the Clarke subdifferential and the upper semicontinuity of $(e_\lambda f)^o(\cdot;\cdot)$ that

$$\begin{aligned}
\limsup_{k\to\infty}\left\langle x_k^*, \frac{\|v-u\|}{\|w-x(t)\|}(w-x_k)\right\rangle &\leqslant \limsup_{k\to\infty}(e_\lambda f)^o\left(x_k; \frac{\|v-u\|}{\|w-x(t)\|}(w-x_k)\right)\\
&\leqslant (e_\lambda f)^o\big(x(t); v-u\big).
\end{aligned}$$

Hence by the equality between the second and third members of (3.11) of Theorem 3.2 and by the derivability at $t \in T$ of the function $r \mapsto e_\lambda f(u + r(v-u))$, we get

$$\limsup_{k\to\infty}\left\langle x_k^*, \frac{\|v-u\|}{\|w-x(t)\|}(w-x_k)\right\rangle \leqslant (e_\lambda f)^o\big(x(t); v-u\big) = (e_\lambda f)'\big(x(t); v-u\big),$$

which translates the condition (iii) of Definition 2.1. The proof of the theorem is then finished. □

Below we use Theorem 3.7 to investigate the Moreau envelope $e_\lambda f$ through selections of the set-valued mapping $E_{\lambda,Q} f(\cdot)$. For this reason we need a result recently obtained in [34, Theorem (c)].

Proposition 3.4. *Let U be a nonempty open convex subset of a Banach space X, Q be a dense subset of U and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a ∂-eds function on U relative to* Dom ∂f*, where ∂ is a presubdifferential included in the Clarke subdifferential. Then if the set-valued mapping ∂f admits a selection on Q which is also a continuous mapping on U, then f is strictly Fréchet differentiable at every point $x \in U$ with $D_F f(x) = \sigma(x)$, so it is of class $\mathcal{C}^1$ on U.*

As a direct consequence of (3.13), Theorem 3.7, and Proposition 3.4 we get the following proposition:

Proposition 3.5. *Let $(X, \|\cdot\|)$ be a Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function. Suppose that f is bounded from below by a quadratic function, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). Let a real $\lambda \in]0, \frac{1}{2\alpha}[$ be given and a dense subset $Q \subset X$. Assume that there exists a continuous mapping $p^* : X \to X^*$ such that:*

$$p^*(q) \in E_{\lambda,X} f(q), \quad \forall q \in Q. \tag{3.27}$$

Then $e_\lambda f(\cdot)$ is strictly Fréchet differentiable at every point $x \in X$ with $D_F f(x) = p^*(x)$, so $e_\lambda f(\cdot)$ is of class $\mathcal{C}^1$ on X.

Of course a good candidate for $p^*(\cdot)$ is any selection of $\frac{1}{\lambda}(\cdot - P_\lambda(\cdot))$, whenever $X = \mathbb{R}^n$. Several facts on this problem can be found in [58] where properties of f and $e_\lambda f$ are investigated with the use of $P_\lambda f(\cdot)$ being single-valued. This problem becomes more difficult, if X is not a finite-dimensional space; we refer to [6] for some facts in Hilbert spaces. Below we give sufficient conditions for finding selections.

Theorem 3.8. *Let $(X, \|\cdot\|)$ be a reflexive Banach space whose norm is uniformly Gâteaux differentiable. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper weakly lower sequentially semicontinuous function bounded from below by a quadratic function. Assume that for some $\lambda \in]0, \frac{1}{2\alpha}[$ the set $P_\lambda f(x)$ reduces to exactly one element for every $x \in X$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). Then*

$$E_{\lambda,X} f(x) = \left\{\frac{1}{\lambda} J\big(x - P_\lambda f(x)\big)\right\}$$

for every $x \in X$. Moreover $e_\lambda f(\cdot)$ is strictly Hadamard differentiable with $D_H e_\lambda f$ continuous from X endowed with the norm topology into X^* endowed with the weak* topology.

Additionally, the Kadec–Klee property of the norm $\|\cdot\|$ gives the continuity of the mapping $P_\lambda f(\cdot)$ (with respect to the norm topology of X), while the Kadec–Klee property of the dual norm $\|\cdot\|_*$ on the dual space X^* gives the strict Fréchet differentiability of $e_\lambda f(\cdot)$, and hence its C^1-smoothness on X.

Proof. Let us fix $x \in X$. It follows from Proposition 3.3 that the set $E_{\lambda,X} f(x)$ is nonempty and weak* compact. Fix any $y^* \in E_{\lambda,X} f(x)$ and any $h \in X$. It follows from the definition of $E_{\lambda,X} f(x)$ that we can find (see (3.15)) sequences $\{y_i^*\}_{i\in\mathbb{N}}$ in X^*, $\{y_i\}_{i\in\mathbb{N}}$ and $\{z_i\}_{i\in\mathbb{N}}$ in X, and $\{\varepsilon_i\}_{i\in\mathbb{N}}$ in $]0,1[$, $\varepsilon_i \to 0$ such that $\langle y_i^*, h\rangle \to \langle y^*, h\rangle$ and such that for all $i \in \mathbb{N}$

$$
\begin{gathered}
e_\lambda f(x) + \varepsilon_i \geqslant f(z_i) + \frac{1}{2\lambda}\|z_i - y_i\|^2,\\
\|y_i - x\| \leqslant \varepsilon_i \quad \text{and} \quad y_i^* \in \partial\big((2\lambda)^{-1}\|\cdot\|^2\big)(y_i - z_i).
\end{gathered}
$$

Of course $y_i^* = \lambda^{-1} J(y_i - z_i)$. Using Proposition 3.3 and the Eberlein–Smulian Theorem, see [29], we infer that there is a subsequence $\{y_{i_k} - z_{i_k}\}_{k\in\mathbb{N}}$ which converges weakly to $x - z$ for some $z \in X$, hence $\{z_{i_k}\}_{k\in\mathbb{N}}$ converges weakly to z. Since

$$
\begin{aligned}
e_\lambda f(x) &\geqslant \limsup_{k\to\infty} f(z_{i_k}) + \frac{1}{2\lambda}\|z_{i_k} - y_{i_k}\|^2\\
&\geqslant \liminf_{k\to\infty} f(z_{i_k}) + \frac{1}{2\lambda}\|z_{i_k} - y_{i_k}\|^2 \geqslant f(z) + \frac{1}{2\lambda}\|z - x\|^2,
\end{aligned}
$$

we get $z = P_\lambda f(x)$ and

$$
\lim_{k\to\infty} f(z_{i_k}) + \frac{1}{2\lambda}\|z_{i_k} - y_{i_k}\|^2 = f(z) + \frac{1}{2\lambda}\|z - x\|^2.
$$

The latter equality and the weak sequential lower semicontinuity of f and $\|\cdot\|^2$ yield $\|y_{i_k} - z_{i_k}\|^2 \to \|x - P_\lambda f(x)\|^2$. Thus the Eberlein–Smulian Theorem (keep in mind that the weak and weak* topology coincide on the dual space X^* of the reflexive space X) and Theorem 3.6 give a subsequence of $\{J(y_i - z_i)\}_{i\in\mathbb{N}}$ converging weakly* to $J(x - P_\lambda f(x))$, hence the whole sequence $\{J(y_i - z_i)\}_{i\in\mathbb{N}}$ converges weakly* to $J(x - P_\lambda f(x))$. Therefore $\langle y^*, h\rangle = \lambda^{-1}\langle J(x - P_\lambda f(x)), h\rangle$ and this being true for every $h \in X$ we get $y^* = \lambda^{-1} J(x - P_\lambda f(x))$, which proves the equality concerning $E_{\lambda,X} f(x)$. Note also that the arguments above ensure that the whole sequence $\{z_i\}_{i\in\mathbb{N}}$ converges weakly to $P_\lambda f(x)$.

In order to get the continuity of the strict Hadamard derivative of $e_\lambda f$ let us fix a sequence $\{x_i\}_{i\in\mathbb{N}} \subset X$ converging to x. Taking x_i and $P_\lambda f(x_i)$, instead of y_i and z_i, respectively, by the above reasoning we get $2(x_i - P_\lambda f(x_i)) \xrightarrow{w} 2(x - P_\lambda f(x))$ and $2\|(x_i - P_\lambda f(x_i))\| \to 2\|(x - P_\lambda f(x))\|$. Thus by Theorem 3.6

$$
D_H e_\lambda f(x_i) = \frac{1}{\lambda} J\big(x_i - P_\lambda f(x_i)\big) \xrightarrow{w^*} \frac{1}{\lambda} J\big(x - P_\lambda f(x)\big),
$$

which implies the norm to weak* continuity of $D_H e_\lambda f(\cdot)$.

Suppose that the norm $\|\cdot\|$ of X satisfies the Kadec–Klee property. In order to get the continuity of $P_\lambda f$, let us take any sequence $\{x_i\}_{i\in\mathbb{N}}$ in X converging to x. Taking x_i and $P_\lambda f(x_i)$, instead of y_i and z_i, respectively, by what precedes we get $2(x_i - P_\lambda f(x_i)) \xrightarrow{w} 2(x - P_\lambda f(x))$ and $2\|(x_i - P_\lambda f(x_i))\| \to 2\|(x - P_\lambda f(x))\|$, which by the Kadec–Klee property of the norm $\|\cdot\|$ gives $2(x_i - P_\lambda f(x_i)) \to 2(x - P_\lambda f(x))$, proving the continuity of $P_\lambda f$ with respect to the norm topology of X.

Let us now assume that the norm $\|\cdot\|_*$ of X^* has the Kadec–Klee property. Take sequences $\{x_i\}_{i\in\mathbb{N}}$, $\{h_i\}_{i\in\mathbb{N}}$, $\{a_i\}_{i\in\mathbb{N}}$ and $\{b_i\}_{i\in\mathbb{N}}$ in X, such that $x_i \to x$, $h_i \to 0$ and $e_\lambda f(x_i) + \|h_i\|^2 \geqslant f(a_i) + \frac{1}{2\lambda}\|x_i - a_i\|^2$, $e_\lambda f(x_i + h_i) + \|h_i\|^2 \geqslant f(b_i) + \frac{1}{2\lambda}\|x_i + h_i - b_i\|^2$ for every $i \in \mathbb{N}$. Observe that from the above reasoning we have $\|x_i - b_i\| \to \|x - P_\lambda f(x)\|$, $x_i - b_i \xrightarrow{w} x - P_\lambda f(x)$ and $\|x_i + h_i - a_i\| \to \|x - P_\lambda f(x)\|$, $x_i + h_i - a_i \xrightarrow{w} x - P_\lambda f(x)$. So Theorem 3.6 ensures $J(x_i - b_i) \xrightarrow{w} J(x - P_\lambda f(x))$, $J(x_i + h_i - a_i) \xrightarrow{w} J(x - P_\lambda f(x))$ and by the Kadec–Klee property of the norm $\|\cdot\|_*$ of X^* we obtain $J(x_i - b_i) \to J(x - P_\lambda f(x))$, $J(x_i + h_i - a_i) \to J(x - P_\lambda f(x))$ (keep in mind that the weak and weak* topology of X^* coincide by the reflexivity of X). We have also

$$
\begin{aligned}
-\|h_i\| + \left\langle \frac{1}{\lambda\|h_i\|} J(x_i - b_i), h_i \right\rangle &\leqslant \frac{-\|h_i\|^2 + \frac{1}{2\lambda}\|x_i + h_i - b_i\|^2 - \frac{1}{2\lambda}\|x_i - b_i\|^2}{\|h_i\|} \\
&\leqslant \frac{e_\lambda f(x_i + h_i) - e_\lambda f(x_i)}{\|h_i\|} \\
&\leqslant \frac{\|h_i\|^2 + \frac{1}{2\lambda}\|x_i + h_i - a_i\|^2 - \frac{1}{2\lambda}\|x_i - a_i\|^2}{\|h_i\|} \\
&\leqslant \|h_i\| + \left\langle \frac{1}{\lambda\|h_i\|} J(x_i + h_i - a_i), h_i \right\rangle,
\end{aligned}
$$

which implies

$$
\lim_{\|h\|\downarrow 0,\, x' \to x} \frac{e_\lambda f(x' + h) - e_\lambda f(x') - \langle D_H e_\lambda f(x), h\rangle}{\|h\|} = 0,
$$

hence the strict Fréchet differentiability follows. Finally, it is easy to see that this strict Fréchet differentiability of the function $e_\lambda f$ at each point of X entails its C^1 property on the space X. □

Corollary 3.3. *Let H be a Hilbert space and $f : H \to \mathbb{R} \cup \{+\infty\}$ be a proper weakly lower semicontinuous function bounded from below by a quadratic function. Assume that for some $\lambda \in \left]0, \frac{1}{2\alpha}\right[$ the set $P_\lambda f(x)$ reduces to exactly one element for every $x \in H$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). Then*

$$
E_{\lambda,X} f(x) = \left\{ \frac{1}{\lambda}\big(x - P_\lambda f(x)\big) \right\}
$$

for every $x \in H$. Moreover the mappings $E_{\lambda,X} f(\cdot)$ and $P_\lambda f(\cdot)$ are continuous (with respect to the norm topology of X) and $e_\lambda f(\cdot)$ is of class C^1.

Proof. The duality mapping being in this case the identity mapping on X, we obtain from the proposition above that $E_{\lambda,X} f(x) = \frac{1}{\lambda}(x - P_\lambda f(x))$, and this combined with the statement of the proposition above justifies the corollary since $\nabla_H e_\lambda f(x) = \frac{1}{\lambda}(x - P_\lambda f(x))$ by (3.13) of Theorem 3.2. □

The following proposition which has its own interest prepares the next theorem where the weak lower semicontinuity of f involved in Theorem 3.8 is replaced by an assumption of continuity of $P_\lambda f$.

Proposition 3.6. *Let $(X, \|\cdot\|)$ be a Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function bounded from below by a quadratic function and $x \in X$ be given. Assume that for some $\lambda \in]0, \frac{1}{2\alpha}[$ the set $P_\lambda f(x)$ is not empty, where α is as in relation* (3.5) (*with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$*)*, and let $z \in P_\lambda f(x)$, $h \in X$ be such that there are $\{t_i\}_{i\in\mathbb{N}}$ in $]0, \infty[$, $\{\varepsilon_i\}_{i\in\mathbb{N}}$ in $[0, \infty[$ with $t_i \downarrow 0$, $\varepsilon_i \downarrow 0$, $\{h_i\}_{i\in\mathbb{N}}$ in X, $h_i \to h$ and $\{z_i\}_{i\in\mathbb{N}}$ in $\operatorname{dom} f$ satisfying*:

(i) $\|x + t_i h_i - z_i\| \to \|x - z\|$ *and* $\varepsilon_i t_i + e_\lambda f(x + t_i h_i) \geqslant f(z_i) + \frac{1}{2\lambda}\|x + t_i h_i - z_i\|^2$, *and*
(ii) $\liminf_{i\to\infty}\langle J(x - z), x - z_i\rangle \geqslant \|x - z\|^2$.

Then

$$\lim_{t\downarrow 0} \frac{e_\lambda f(x + th) - e_\lambda f(x)}{t} = \langle \lambda^{-1} J(x - z), h\rangle.$$

Moreover, if (i) *and* (ii) *hold true for every $h \in X$ then $e_\lambda f$ is Hadamard differentiable and $D_H e_\lambda f(x) = \lambda^{-1} J(x - z)$.*

Proof. Let $z \in P_\lambda f(x)$ be as in the assumption. Fix $h \in X$, $\{t_i\}_{i\in\mathbb{N}}$ and $\{\varepsilon_i\}_{i\in\mathbb{N}}$ in $]0, \infty[$ with $t_i \downarrow 0$, $\varepsilon_i \downarrow 0$, $\{h_i\}_{i\in\mathbb{N}}$ in X with $h_i \to h$, and $\{z_i\}_{i\in\mathbb{N}}$ in X such that:

- $\|x + t_i h_i - z_i\| \to \|x - z\|$ and $\varepsilon_i t_i + e_\lambda f(x + t_i h_i) \geqslant f(z_i) + \frac{1}{2\lambda}\|x + t_i h_i - z_i\|^2$, and
- $\liminf_{i\to\infty}\langle J(x - z), x - z_i\rangle \geqslant \|x - z\|^2$.

By (3.12) of Theorem 3.2 we get

$$\begin{aligned} d^- e_\lambda(x; h) &= \lim_{t\downarrow 0} \frac{e_\lambda f(x + th) - e_\lambda f(x)}{t} = \lim_{i\to\infty} \frac{e_\lambda f(x + t_i h_i) - e_\lambda f(x)}{t_i} \\ &\geqslant \lim_{i\to\infty} \frac{-\varepsilon_i t_i + f(z_i) + \frac{1}{2\lambda}\|x + t_i h_i - z_i\|^2 - f(z_i) - \frac{1}{2\lambda}\|x - z_i\|^2}{t_i} \\ &= \lim_{i\to\infty} \frac{-\varepsilon_i t_i + \frac{1}{2\lambda}\|x + t_i h_i - z_i\|^2 - \frac{1}{2\lambda}\|x - z_i\|^2}{t_i} \geqslant \frac{1}{\lambda} \lim_{i\to\infty}\langle J(x - z_i), h\rangle. \end{aligned}$$

Further, from Theorem 3.6 we have $J(x - z_i) \xrightarrow{w^*} J(x - z)$, hence for every $h \in X$ we obtain

$$d^-(e_\lambda f)(x; h) \geqslant \frac{1}{\lambda}\langle J(x - z), h\rangle. \tag{3.28}$$

On the other hand we know by (3.11) of Theorem 3.2 that the limit $\lim_{t\uparrow 0} \frac{e_\lambda f(x-th)-e_\lambda f(x)}{t}$ exists, thus

$$\lim_{t\downarrow 0} \frac{e_\lambda f(x)-e_\lambda f(x+th)}{t} \geqslant \lim_{t\downarrow 0} \frac{f(z)+\frac{1}{2\lambda}\|x-z\|^2 - f(z) - \frac{1}{2\lambda}\|x+th-z\|^2}{t}$$
$$= \frac{1}{\lambda}\langle J(x-z), -h\rangle.$$

So by (3.11) of Theorem 3.2 and (3.28) we get

$$\frac{1}{\lambda}\langle J(x-z), h\rangle \geqslant d^-(e_\lambda f)(x;h) \geqslant \frac{1}{\lambda}\langle J(x-z), h\rangle,$$

which implies by (3.12) of Theorem 3.2 the desired equality

$$\lim_{t\downarrow 0} \frac{e_\lambda f(x+th)-e_\lambda f(x)}{t} = \langle \lambda^{-1} J(x-z), h\rangle.$$

Moreover $\frac{1}{\lambda}J(x-z) \in \partial^- e_\lambda f(x)$, whenever (i) and (ii) hold true for every $h \in X$, hence the final statement of the proposition follows from Theorem 3.5. □

Theorem 3.9. *Let $(X, \|\cdot\|)$ be a Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function bounded from below by a quadratic function and $x \in X$ be given. Let $\lambda \in]0, \frac{1}{2\alpha}[$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). Assume that $P_\lambda f(\cdot)$ is single-valued on an open set U containing x.*

(a) *If $P_\lambda f(\cdot)$ is directionally continuous at x, that is, the mappings $t \mapsto P_\lambda f(x+th)$ are continuous at $0 \in \mathbb{R}$ for every $h \in X$, then $e_\lambda f$ is strictly Hadamard differentiable at x and $D_H e_\lambda f(x) = \lambda^{-1} J(x - P_\lambda f(x))$.*
(b) *If $P_\lambda f(\cdot)$ is continuous on U and the dual norm $\|\cdot\|_*$ of X^* has the Kadec–Klee property, then $e_\lambda f$ is of class C^1 on U.*

Proof. For any fixed $h \in X$ select $t_i \downarrow 0$ such that $x + t_i h \in U$ for all $i \in \mathbb{N}$, and put $z_i := P_\lambda f(x + t_i h)$ for all $i \in \mathbb{N}$, and $z := P_\lambda f(x)$. The directional continuity assumption of $P_\lambda f(\cdot)$ at x ensures $z_i \to z$ as $i \to \infty$ hence conditions (i) and (ii) of Proposition 3.6 are fulfilled with $\varepsilon_i = 0$ and $h_i = h$ for all i, so the assertion (a) of the theorem follows from Proposition 3.6.

If in addition the dual norm $\|\cdot\|_*$ fulfills the Kadec–Klee property, we know by (3.24) that J is norm–norm continuous, so the C^1 property of $e_\lambda f$ is a consequence of the formula $D_H e_\lambda f(x) = \lambda^{-1} J(x - P_\lambda f(x))$ in (a). □

4. Differentiability of the distance function

In this section we apply the above results to get (sub)differentiability properties of the distance function. Let us start with the observation, in addition to Remark 3.3, that another direct consequence of Theorem 3.2 is the following result concerning some directional differential properties of the distance function. This property was first established in [8, Theorem 8].

Theorem 4.1. *Let $(X, \|\cdot\|)$ be a normed vector space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and let $S \subset X$ be a nonempty closed subset. Then for all $x \notin S$ and $h \in X$ the following hold*:

$$-d_S'(x; h) = d^-(-d_S)(x; h) = d_S^o(x; -h) = (-d_S)^o(x; h), \tag{4.1}$$

that is, $-d_S$ is directionally subregular at any point outside S.

Proof. Fix $x \notin S$ and $h \in X$. We have, see Remark 3.1, $e_{\frac{1}{2}}\psi_S(x) = d_S^2(x)$ and by equalities (2.15)

$$(e_{\frac{1}{2}}\psi_S)^o(x; h) = 2d_S(x)d_S^o(x; h) \quad \text{and} \quad d^-(-e_{\frac{1}{2}}\psi_S)(x; h) = 2d_S(x)d^-(-d_S)(x; h). \tag{4.2}$$

We know that $(e_{\frac{1}{2}}\psi_S)^o(x; -h) = (-e_{\frac{1}{2}}\psi_S)^o(x; h)$, see (2.9). It follows from (3.11) and (3.12) in Theorem 3.2 that

$$d^-(-e_{\frac{1}{2}}\psi_S)(x; h) = (e_{\frac{1}{2}}\psi_S)^o(x; -h),$$

so, by (4.2) we get

$$d^-(-d_S)(x; h) = d_S^o(x; -h) = (-d_S)^o(x; h),$$

which confirms the directional subregularity of the function $-d_S$. $\square$

An immediate consequence of (4.1) and the upper semicontinuity of $d_S^o(\cdot,\cdot)$, see also Remark 3.3, is the following property of the directional derivative of the distance function.

Corollary 4.1. *Let $(X, \|\cdot\|)$ be a normed vector space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and let $S \subset X$ be a nonempty closed subset. Then the function f defined on $X \setminus S$ by*

$$f(x) := \sup_{h \in \mathbb{B}_X} d_S'(x; h)$$

is lower semicontinuous.

Corollary 4.2. *Let $(X, \|\cdot\|)$ be a normed vector space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and let $S \subset X$ be a nonempty closed subset. Assume that there exist a function f bounded from below by a quadratic function and a positive $\lambda \in]0, \frac{1}{2\alpha}[$ such that $e_\lambda f(\cdot) = -d_S^2(\cdot)$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). Then d_S^2 is strictly Hadamard differentiable on X.*

Proof. It follows from Theorem 4.1 and Remark 3.3 that both functions $-d_S^2(\cdot)$ and $d_S^2(\cdot) = -e_\lambda f(\cdot)$ are Clarke directionally subregular. So for every $x \in X$ both functions $d^-(d_S^2)(x; \cdot)$ and $-d^-(d_S^2)(x; \cdot)$ are convex and continuous, thus they are linear and $d^-(d_S^2)(x; \cdot) = (d_S^2)^o(x; \cdot)$, which gives the strict Hadamard differentiability of d_S^2. $\square$

Using (2.15), Theorem 3.5 and Corollary 3.1, as simple consequence we get the following characterizations of the strict Hadamard differentiability of the distance function.

Theorem 4.2. *Let $(X, \|\cdot\|)$ be a normed space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and let $S \subset X$ be a nonempty closed set and $x \in X \setminus S$. Then, the following assertions are equivalent*:

(a) *d_S is Clarke directionally subregular at x.*
(b) *$\partial^- d_S(x) \neq \emptyset$.*
(c) *$\partial_C d_S(x)$ is a singleton.*
(d) *$\partial^- d_S(x) = \partial_C d_S(x)$ and the set $\partial_C d_S(x)$ is singleton.*
(e) *d_S is strictly Hadamard differentiable at x.*
(f) *d_S is Gâteaux differentiable at x.*

Suppose in addition that the space $(X, \|\cdot\|)$ is reflexive, and the dual norm $\|\cdot\|_*$ of the dual space X^* has the Kadec–Klee property, and $P_S(x) \neq \emptyset$. Then each of the above conditions is equivalent to each of the following properties:

(g) *d_S is Fréchet differentiable at x.*
(h) *d_S is strictly Fréchet differentiable at x.*

Proof. Let f be the indicator function of the set S, that is, $f = \psi_S$. Then for all $\lambda > 0$

$$e_\lambda f(u) = \frac{1}{2\lambda} d_S^2(u), \quad \forall u \in X.$$

Since f satisfies all the assumptions of Theorem 3.2, so the equivalences (a) $\Leftrightarrow$ (b) $\Leftrightarrow \cdots \Leftrightarrow$ (f) are simple consequences of Theorem 3.5, Corollary 3.1.

To prove the implication (f) $\Rightarrow$ (h) under the additional assumptions of the theorem, it suffices to consider the concerned properties with d_S^2 in place of d_S. So, suppose that the space X is reflexive and the dual norm $\|\cdot\|_*$ of the dual space X^* has the Kadec–Klee property and $P_S(x) \neq \emptyset$. Take any $z \in P_S(x)$, by Proposition 3.1 we have $D_G d_S^2(x) = 2J(x - z)$. First observe that it follows from (3.11) and from item (e) that

$$E_{\frac{1}{2},X} \psi_S(x) = \left\{ D_G e_{\frac{1}{2}} \psi_S(x) \right\} = \left\{ D_G d_S^2(x) \right\}. \tag{4.3}$$

Consider any sequence $\{z_i\}_{i \in \mathbb{N}}$ in S such that $d_S^2(x) = \lim_{i \to \infty} \|x - z_i\|^2$. The equality (4.3) and the definition of $E_{\frac{1}{2},X} \psi_S(x)$ yield $2J(x - z_i) \xrightarrow{w} D_G d_S^2(x)$ (in the weak topology, we use the reflexivity of the space and instead of the weak* topology we use the weak topology in X^*). Thus we have $2J(x - z_i) \xrightarrow{w} D_G d_S^2(x)$ and

$$\begin{aligned} \left\| D_G d_S^2(x) \right\|_* &= 2 \left\| J(x - z) \right\|_* = 2\|x - z\| = 2 d_S(x) \\ &= 2 \lim_{i \to \infty} \|x - z_i\| = 2 \lim_{i \to \infty} \left\| J(x - z_i) \right\|_*, \end{aligned}$$

so by the Kadec–Klee property of $\|\cdot\|_*$ we have the convergence with respect to the norm topology, that is,

$$2J(x - z_i) \to D_G d_S^2(x). \tag{4.4}$$

Now take sequences $\{x_i\}_{i\in\mathbb{N}}$ and $\{h_i\}_{i\in\mathbb{N}}$ in X with $\|h_i\| \neq 0$, $\{a_i\}_{i\in\mathbb{N}}$ in S and $\{b_i\}_{i\in\mathbb{N}}$ in S, such that $x_i \to x$, $h_i \to 0$ and

$$d_S^2(x_i) + \|h_i\|^2 \geqslant \|x_i - a_i\|^2 \geqslant d_S^2(x_i),$$
$$d_S^2(x_i + h_i) + \|h_i\|^2 \geqslant \|x_i + h_i - b_i\|^2 \geqslant d_S^2(x_i + h_i)$$

for every $i \in \mathbb{N}$. On the one hand, the latter inequalities imply

$$d_S^2(x) = \lim_{i\to\infty} \|x_i + h_i - a_i\|^2 \quad \text{and} \quad d_S^2(x) = \lim_{i\to\infty} \|x_i - b_i\|^2.$$

Combining this with (4.4), we obtain that $\{2J(x - a_i)\}_{i\in\mathbb{N}}$ and $\{2J(x - b_i)\}_{i\in\mathbb{N}}$ both converge in norm to $D_G d_S^2(x)$, so the norm–norm continuity of J (see (3.24)) yields

$$\lim_{i\to\infty} \left\| D_G d_S^2(x) - 2J(x_i + h_i - a_i) \right\|_* = \lim_{i\to\infty} \left\| D_G d_S^2(x) - 2J(x_i - b_i) \right\|_* = 0. \tag{4.5}$$

On the other hand, the same inequalities also guarantee that

$$\begin{aligned} -\|h_i\| + \frac{\langle 2J(x_i - b_i), h_i\rangle}{\|h_i\|} &\leqslant \frac{-\|h_i\|^2 + \|x_i + h_i - b_i\|^2 - \|x_i - b_i\|^2}{\|h_i\|} \\ &\leqslant \frac{d_S^2(x_i + h_i) - d_S^2(x_i)}{\|h_i\|} \leqslant \frac{\|h_i\|^2 + \|x_i + h_i - a_i\|^2 - \|x_i - a_i\|^2}{\|h_i\|} \\ &\leqslant \|h_i\| + \frac{\langle 2J(x_i + h_i - a_i), h_i\rangle}{\|h_i\|}. \end{aligned}$$

According to (4.5) we deduce

$$\lim_{\|h\|\downarrow 0,\, x'\to x} \frac{d_S^2(x' + h) - d_S^2(x') - \langle D_G d_S^2(x), h\rangle}{\|h\|} = 0,$$

which translates the property (h). $\square$

Several results on the differentiability of the distance function can be found in papers [8,12,19, 24–27]. For results on smallness of sets of points where the distance function is not differentiable (assuming differentiability conditions of the norm), we refer to [39,63]. In [24, Corollary 3.6] the equivalence between the Gâteaux (with an additional condition that $\|D_G d_S(x)\| = 1$) and Fréchet differentiability was established under the assumptions that the norm $\|\cdot\|$ of X is both Fréchet differentiable and uniformly Gâteaux differentiable, and the norm of X^* is Fréchet differentiable. In order to compare the equivalence between items (f) and (g) from Theorem 4.2 with that from [24, Corollary 3.6] we note that the assumption of Fréchet differentiability of the norm $\|\cdot\|$ of X yields the Kadec–Klee property of the dual norm $\|\cdot\|_*$ of X^* whenever X is a reflexive Banach space, see [18, Theorem 1, (i)(iv), p. 22 and Lemma 1, p. 29]; note also that the assumption of Fréchet differentiability of the dual norm $\|\cdot\|_*$ of X^* implies the reflexivity of X, see [18, Corollary 1, p. 34]. Let us also observe that item (d) from Theorem 4.2 and Proposition 3.1, Remark 3.1 imply that $\|D_G d_S(x)\| = 1$, whenever the Gâteaux derivative $D_G d_S(x)$ exists.

Below we will see that one of consequences of Theorem 4.2 is that the nonemptiness of $P_S(x)$ forces the differentiability of $d_S^2(\cdot)$ at every point of the segment $[z,x[$, whenever X is a Hilbert space and $z \in P_S(x)$. In fact, the result will follow from the next proposition (which is a consequence of Proposition 3.6). This proposition says that for any x outside a closed set S of a Hilbert space H with $P_S(x) \neq \emptyset$ and for $z \in P_S(x)$ one has $2\rho(x-z) \in \partial^- d_S^2(\rho x + (1-\rho)z)$ for each $\rho \in [0,1[$. Its complements in some sense the result that $\partial^- d_S^2(u) = \{2(u - P_S(u))\}$, for any u outside S, whenever the Hilbert space H is finite-dimensional and $P_S(u)$ is a singleton (see [51, Example 8.53]).

Proposition 4.1. *Let H be a Hilbert space and $S \subset H$ be a nonempty subset and $x \in H$ be given such that $P_S(x) \neq \emptyset$. Then for every $z \in P_S(x)$, $\rho \in [0,1[$ we have $2\rho(x-z) \in \partial^- d_S^2(\rho x + (1-\rho)z)$.*

Proof. Fix $z \in P_S(x)$, $\rho \in [0,1[$ and $h \in H$. Of course $z \in P_S(\rho x + (1-\rho)z)$. For every $\mu > 0$ let us put

$$\begin{aligned}\delta(\mu) := \inf\bigl\{&\|z-s\|^2:\ s \in S \text{ and}\\ &\bigl\|\rho x + (1-\rho)z + th - s\bigr\|^2 \leqslant d_S^2\bigl(\rho x + (1-\rho)z + th\bigr) + t\mu,\ \text{and } t \in]0,\mu]\bigr\} \quad \text{and}\\ \delta :={}& \lim_{\mu\downarrow 0} \delta(\mu).\end{aligned}$$

If $\delta = 0$ then by Proposition 3.6 we obtain

$$\lim_{t\downarrow 0} \frac{d_S^2(\rho x + (1-\rho)z + th) - d_S^2(\rho x + (1-\rho)z)}{t} = \bigl\langle 2\rho(x-z), h\bigr\rangle. \tag{4.6}$$

Suppose $\delta > 0$. Take $\{t_i\}_{i\in\mathbb{N}}$ in $]0,1[$, $t_i \downarrow 0$. If $\|\rho x + (1-\rho)z + t_i h - z\|^2 = d_S^2(\rho x + (1-\rho)z + t_i h)$ we set $s_i := z$, and if not, that is,

$$\bigl\|\rho x + (1-\rho)z + t_i h - z\bigr\|^2 > d_S^2\bigl(\rho x + (1-\rho)z + t_i h\bigr)$$

we can take $s_i \in S$ such that

$$\begin{aligned}&\bigl\|\rho x + (1-\rho)z + t_i h - s_i\bigr\|^2\\ &\quad < \min\bigl\{\bigl\|\rho x + (1-\rho)z + t_i h - z\bigr\|^2, d_S^2\bigl(\rho x + (1-\rho)z + t_i h\bigr) + t_i^2\bigr\}.\end{aligned}$$

So the sequence $\{s_i\}_{i\in\mathbb{N}}$ in S satisfies

$$\bigl\|\rho x + (1-\rho)z + t_i h - s_i\bigr\|^2 \leqslant d_S^2\bigl(\rho x + (1-\rho)z + t_i h\bigr) + t_i^2$$

and

$$\bigl\|\rho x + (1-\rho)z + t_i h - z\bigr\|^2 \geqslant \bigl\|\rho x + (1-\rho)z + t_i h - s_i\bigr\|^2$$

for every $i \in \mathbb{N}$. Observe that for every $i \in \mathbb{N}$ we have $\|x - s_i\|^2 \geqslant \|x - z\|^2$ so $2\langle x - z, z - s_i\rangle + \|z - s_i\|^2 \geqslant 0$, and since

$$\left\|\rho x + (1-\rho)z + t_i h - z\right\|^2 \geqslant \left\|\rho x + (1-\rho)z + t_i h - s_i\right\|^2$$

we obtain

$$\begin{aligned}0 &\geqslant 2\left\langle \rho x + (1-\rho)z + t_i h - z, z - s_i\right\rangle + \|z - s_i\|^2\\ &\geqslant 2\langle t_i h, z - s_i\rangle + (1-\rho)\|z - s_i\|^2 \geqslant 2\langle t_i h, z - s_i\rangle + (1-\rho)\delta(t_i),\end{aligned}$$

hence $0 \geqslant 1-\rho$, since $\lim_{\mu\downarrow 0}\delta(\mu) = \delta > 0$, which is a contradiction. Consequently $\delta = 0$ and by (4.6) we get the following equality

$$\lim_{t\downarrow 0}\frac{d_S^2(\rho x + (1-\rho)z + th) - d_S^2(\rho x + (1-\rho)z)}{t} = \left\langle 2\rho(x-z), h\right\rangle$$

and the statement of the proposition is valid. □

Now we can state an immediate consequence of the proposition above and Theorem 4.2 ((b) and (g)).

Corollary 4.3. *Let H be a Hilbert space and $S \subset H$ be a nonempty closed subset such that*

$$H \setminus S \subset \bigcup_{x\in(\mathrm{Dom}\, P_S)\setminus S,\ z\in P_S(x)} [z, x[,$$

where $\mathrm{Dom}\, P_S := \{u \in X\colon\ P_S(u) \neq \emptyset\}$. *Then $\partial^- d_S^2(x) \neq \emptyset$ for every $x \in H \setminus S$, thus $d_S^2(\cdot)$ is continuously Fréchet differentiable on H, that is, the Fréchet derivative is continuous.*

Observe that for every closed convex subset $S \subset X$ we have the Clarke directional subregularity of d_S^2 at every $x \in X$. Thus we have the following second corollary of Theorem 4.2.

Corollary 4.4. *Let $(X, \|\cdot\|)$ be a normed space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and let $S \subset X$ be a nonempty closed convex set. Then the following assertions hold true*:

(a) *For all $x \in X$ the set $\partial_C d_S^2(x)$ is singleton.*
(b) *For all $x \in X$ the function d_S^2 is strictly Hadamard differentiable at x.*
(c) *For all $x \in X \setminus S$ the function d_S is strictly Hadamard differentiable at x.*

Below we establish that the Vlasov condition, see [57, p. 56] and also [55,56],

$$\limsup_{y\to 0}\frac{d_S(x+y) - d_S(x)}{\|y\|} = 1, \quad \forall x \notin S \tag{4.7}$$

together with the following one

$$\liminf_{\|x\|\to+\infty}\left(d_{[a,b]}(x) - d_S(x)\right) \geqslant 0, \quad \forall a, b \in S,\ a \neq b \tag{4.8}$$

is equivalent to the convexity of S, where $S \subset X$ is a given closed set. We show that this new condition, that is, condition (4.8), is always satisfied whenever the norm of the space is uniformly Gâteaux differentiable (without any additional assumptions on the set S). We must say that L.P. Vlasov showed

$$d_{[a,b]}(x) - d_S(x) \geqslant 0, \quad \forall a, b \in S,\ a \neq b \text{ and } \forall x \in X, \tag{4.9}$$

whenever the dual norm of X^* is strictly convex (or equivalently, rotund) and simultaneously (4.7) is fulfilled, see the proof of Theorem 2 of [56] and [55, Theorem 1, Lemma 1 and Lemma 2]. We also point out that the condition (4.7) is a key in proving the convexity of Tchebyshev sets; this was observed by L.P. Vlasov when the strict convexity of the dual norm $\|\cdot\|_*$ of the dual space X^* is assumed, see [55,56]. We recall that a nonempty set S of a normed space $(X, \|\ \|)$ is a *Tchebyshev* set provided that $P_S(x)$ is a singleton for any $x \in X$, see (2.1) for the definition of $P_S(x)$. Obviously, any such set is closed in $(X, \|\ \|)$.

Theorem 4.3. *Let X be a Banach space and $S \subset X$ be a closed set. Then S is convex if and only if both properties* (i) *and* (ii) *below hold*:

(i)
$$\limsup_{h \to 0} \frac{d_S(x+h) - d_S(x)}{\|h\|} = 1, \quad \forall x \notin S. \tag{4.10}$$

(ii) *For all $a, b \in S$, with $a \neq b$, we have*

$$\liminf_{\|x\| \to +\infty} \big(d_{[a,b]}(x) - d_S(x)\big) \geqslant 0. \tag{4.11}$$

Proof. In order to prove the convexity of S under (i) and (ii) let us take $a', b' \in S$. If the segment $[a', b'] \subset S$ then we are done. Let us consider the case $[a', b'] \setminus S \neq \emptyset$. Take $a, b \in S \cap [a', b']$ such that $]a, b[\cap S = \emptyset$. Put $\varepsilon := \frac{d_S(\frac{a+b}{2})}{2}$ and

$$f(x) := -d_S^2(x) + d_{[a,b]}^2(x) + \varepsilon \left\| x - \frac{a+b}{2} \right\|$$

for every $x \in X$. First note that

$$f(x) = \big(d_{[a,b]}(x) - d_S(x)\big)\big(d_{[a,b]}(x) + d_S(x)\big) + \varepsilon \left\| x - \frac{a+b}{2} \right\|,$$

so

$$\liminf_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} = \liminf_{\|x\| \to +\infty} \big(d_{[a,b]}(x) - d_S(x)\big) \frac{d_{[a,b]}(x) + d_S(x)}{\|x\|} + \varepsilon$$

so the relation (4.11) gives

$$\liminf_{\|x\| \to +\infty} \frac{f(x)}{\|x\|} \geqslant \varepsilon. \tag{4.12}$$

The latter inequality together with $\frac{a+b}{2} \in [a,b]$ ensures that

$$-\infty < \inf_{x\in X} f(x) \leqslant -d_S^2\left(\frac{a+b}{2}\right) = -4\varepsilon^2 < -2\varepsilon^2.$$

Then, we may for each $i \in \mathbb{N}$, apply the Ekeland variational principle to obtain some $x_i \in X$ such that $f(x_i) \to \inf_{x\in X} f(x) < -2\varepsilon^2$ (this forces $x_i \notin S$ for $i \in \mathbb{N}$ large enough, say $i \geqslant i_0$, and, by (4.12), that the sequence $\{x_i\}_{i\in\mathbb{N}}$ is bounded) and such that

$$f(x) + i^{-1}\|x - x_i\| \geqslant f(x_i) \quad \text{for every } x \in X. \tag{4.13}$$

Fix any integer $i \geqslant i_0$. From (4.13) we deduce for all $h \in X$ that

$$\begin{aligned} &i^{-1}\|h\| + \varepsilon\left[\left\|x_i + h - \frac{a+b}{2}\right\| - \left\|x_i - \frac{a+b}{2}\right\|\right] + d_{[a,b]}^2(x_i+h) - d_{[a,b]}^2(x_i) \\ &\quad \geqslant d_S^2(x_i+h) - d_S^2(x_i), \end{aligned}$$

which ensures that

$$i^{-1} + \varepsilon + 2d_{[a,b]}(x_i) \geqslant 2d_S(x_i) \limsup_{h\to 0} \frac{d_S(x_i+h) - d_S(x_i)}{\|h\|}.$$

Using the assumption (i) the latter inequality is equivalent to

$$\frac{(i^{-1}+\varepsilon)}{2} + d_{[a,b]}(x_i) \geqslant d_S(x_i),$$

and hence

$$\frac{(i^{-1}+\varepsilon)^2}{4} + d_{[a,b]}^2(x_i) + \left(i^{-1}+\varepsilon\right)d_{[a,b]}(x_i) \geqslant d_S^2(x_i).$$

Since $f(x_i) \to \inf_{x\in X} f(x) < -2\varepsilon^2$, we have

$$\begin{aligned} -2\varepsilon^2 > \inf_{x\in X} f(x) &= \lim_{i\to\infty}\left(d_{[a,b]}^2(x_i) - d_S^2(x_i) + \varepsilon\left\|x_i - \frac{a+b}{2}\right\|\right) \\ &\geqslant \limsup_{i\to\infty}\left(d_{[a,b]}^2(x_i) - \frac{(i^{-1}+\varepsilon)^2}{4} - d_{[a,b]}^2(x_i) - \left(i^{-1}+\varepsilon\right)d_{[a,b]}(x_i) + \varepsilon\left\|x_i - \frac{a+b}{2}\right\|\right) \\ &\geqslant \limsup_{i\to\infty}\left(-\frac{(i^{-1}+\varepsilon)^2}{4} - \left(i^{-1}+\varepsilon\right)d_{[a,b]}(x_i) + \varepsilon\left\|x_i - \frac{a+b}{2}\right\|\right) \\ &\geqslant \limsup_{i\to\infty}\left(-\frac{(i^{-1}+\varepsilon)^2}{4} - i^{-1}d_{[a,b]}(x_i)\right) = -\frac{\varepsilon^2}{4}, \end{aligned}$$

a contradiction (note that the last equality is due to the boundedness of $\{x_i\}_{i\in\mathbb{N}}$). Hence $[a', b'] \subset S$ for all $a', b' \in S$, that is, the set S is convex.

Conversely suppose that S is convex. It is obvious that condition (ii) is satisfied. Fix any x outside S and choose $x^* \in \partial d_S(x)$. Through the definition of the subdifferential of convex functions, it is easily seen that $\|x^*\|_* = 1$, thus there exists a sequence $\{u_i\}_{i\in\mathbb{N}}$ with $\|u_i\| = 1$ such that $\langle x^*, u_i\rangle \to \|x^*\|_* = 1$. Taking a sequence $\{t_i\}_{i\in\mathbb{N}}$ tending to 0 with $t_i > 0$ and putting $h_i := t_i u_i$, we obtain $\|h_i\| \to 0$ and

$$\frac{d_S(x+h_i) - d_S(x)}{\|h_i\|} \geqslant \langle x^*, u_i\rangle \quad \text{hence} \quad \limsup_{\|h\|\to 0} \frac{d_S(x+h) - d_S(x)}{\|h\|} = 1,$$

which finishes the proof of the theorem. □

In the following proposition it is established that the relation (4.11) holds true in any normed space with uniformly Gâteaux differentiable norm without additional assumptions on the set S.

Proposition 4.2. *Let $(X, \|\cdot\|)$ be a normed space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable and $S \subset X$ be a closed subset with $a, b \in S$ and $a \neq b$. Then one has*

$$\liminf_{\|x\|\to\infty} \big(d_{[a,b]}(x) - d_S(x)\big) \geqslant 0.$$

Proof. Put $f(\cdot) = d_{[a,b]}(\cdot) - d_S(\cdot)$ and take any selection p of $P_{[a,b]}$, that is $p(x) \in P_{[a,b]}(x)$ for every $x \in X$. Let us assume that there is a sequence $\{x_i\}_{i\in\mathbb{N}} \subset X$ with $\|x_i\| \to \infty$ such that

$$\lim_{i\to\infty} f(x_i) < -\mu \tag{4.14}$$

for some $\mu > 0$. We may suppose that for every $i \in \mathbb{N}$, $p(x_i) \neq b$, $x_i \notin [a,b]$ and $f(x_i) < -\mu$ (in fact this is true for i large enough, otherwise $\lim_{i\to\infty} f(x_i) \geqslant 0$). Since $a \in S$ we get

$$\big\|x_i - p(x_i)\big\| + \mu \leqslant \|x_i - a\|$$

and

$$\mu \leqslant \frac{\|y_i + t_i(p(x_i) - a)\| - \|y_i\|}{t_i}$$

for $t_i := \|x_i - p(x_i)\|^{-1}$, $y_i := t_i(x_i - p(x_i))$. Note that $\|y_i\| = 1$. The compactness of the segment $[a,b]$ and the uniform Gâteaux differentiability of the norm yield

$$\mu \leqslant \liminf_{i\to\infty} \big\langle J(y_i), p(x_i) - a\big\rangle,$$

we recall that J stands for the Gâteaux derivative of $\frac{1}{2}\|\cdot\|^2$. Further, the point $p(x_i)$ being a minimizer of $\|x_i - \cdot\|$ over the convex set $[a,b]$, we have $\langle -J(x_i - p(x_i)), u - p(x_i)\rangle \geqslant 0$ for all $u \in [a,b]$. Since $p(x_i) \neq b$, we deduce that $\langle J(y_i), p(x_i) - a\rangle = 0$ (we have $t_i J(x_i - p(x_i)) = J(y_i)$ by (3.4)). So $\mu \leqslant 0$, which is a contradiction, and hence (4.14) does not hold. □

The condition (4.11) may be given in several equivalent forms, for example

$$\liminf_{\|x\|\to+\infty} \frac{d^p_{[a,b]}(x) - d^p_S(x)}{\|x\|^{p-1}} \geqslant 0,$$

with $p \in [1, \infty[$. Moreover the proposition above holds true for any compact subset Q taken instead of $\{a, b\}$ and $\operatorname{co} Q$ instead of $[a, b]$. This allows to rewrite the above result in many different ways. Let us also point out that the property given in the proposition above is closely linked with the differentiability of the norm. In order to illustrate that, we consider the following remark.

Remark 4.1. Now let us remark that dropping the Gâteaux differentiability assumption of the norm, we can loose property (4.11) even in a finite-dimensional case. In order to see that, we use the space $X := \mathbb{R}^2$ endowed with the norm $\|(r, s)\| := |r| + |s|$. Consider $a = (-1, 0)$, $b = (1, 0)$, and $x_i = (0, i)$ for all $i \in \mathbb{N}$, so $\|x_i\| = i$. For $S := \{a, b\}$ the equalities $d_{[a,b]}(x_i) = i$ and $d_S(x_i) = i + 1$ give

$$d_{[a,b]}(x_i) - d_S(x_i) = -1 \quad \text{hence} \quad \liminf_{\|x\| \to +\infty} d_{[a,b]}(x) - d_S(x) \leqslant -1.$$

This makes clear that even in the simple space $\mathbb{R}^2$ the property from the proposition may fail for some usual norms.

The following result was obtained by L.P. Vlasov [56, Theorem 3] in the setting of a Banach space $(X, \|\cdot\|)$ whose dual norm $\|\cdot\|_*$ of X^* was assumed to be strictly convex. Below we provide a new proof of it as a consequence of Theorem 4.3 for X being a Banach space whose norm is uniformly Gâteaux differentiable.

Theorem 4.4. *Let X be a Banach space whose norm is uniformly Gâteaux differentiable and $S \subset X$ be a Tchebyshev set with continuous metric projection. Then S is convex.*

Proof. Fix any x outside S, put $u := (x - P_S(x))/\|x - P_S(x)\|$ and put also $t_i := 1/i$ for every integer $i \geqslant 1$. By Theorem 3.9(a) for each i we have

$$\begin{aligned} 2d_S(x) = 2\langle J\big(x - P_S(x)\big), u\rangle &= \lim_{i\to\infty} \frac{d_S^2(x + t_i u) - d_S^2(x)}{t_i} \\ &= 2d_S(x) \lim_{i\to\infty} \frac{d_S(x + t_i u) - d_S(x)}{\|t_i u\|}, \end{aligned}$$

hence

$$\limsup_{\|h\|\to 0} \frac{d_S(x + h) - d_S(x)}{\|h\|} = 1.$$

The result of the theorem then follows from Theorem 4.3 and Proposition 4.2. $\square$

If X is a Hilbert space and the set $S \subset X$ is weakly closed and Tchebyshev, then $E_{\frac{1}{2},X}\psi_S(\cdot)$ is a single-valued monotone and Lipschitz continuous mapping, so the selection $p^*(\cdot)$ in (3.27) is exactly $E_{\frac{1}{2},X}\psi_S(\cdot)$ and $d_S^2(\cdot)$ is a $C^{1,1}$ function, see [1,36,49]. (We recall that a function is $C^{1,1}$ on an open set U if it is Fréchet differentiable on U and its Fréchet derivative is locally Lipschitz continuous on U.) However if the set S is only Tchebyshev, so it has to be strongly closed, then it is still an open question if it is convex: see [36], where it was stated: “*However, even in Hilbert* *space it remains unknown whether a Tchebyshev set must be convex, or, equivalently, whether it must be weakly closed*"; see [28, Problem 5] *The Possible Convexity of a Tchebyshev Set in a Hilbert Space*, where a recent review of some achievements in solving the problem is given and where it is observed in the Hilbert setting that if the set is Tchebyshev then the Gâteaux differentiability of $d_S^2(\cdot)$ on X is equivalent to the convexity of S. Other important surveys are [5, 7,57] and many facts concerning the convexity of Tchebyshev sets can be found in [16]. We refer to Theorem 4.2 above for a list of equivalent derivative properties characterizing the convexity. We refer also to [50] where a conjecture, whose positive solution would provide an example of a nonconvex Tchebyshev set in an infinite-dimensional real Hilbert space, is posed.

It is rather obvious that for a given set S the existence of continuous selection in (3.27) is linked with the continuity of the projection mapping $P_S(\cdot)$, whenever $f = \psi_S$. There are a lot of results on this subject (guaranteeing the continuity); for old ones, see for example [37], for recent ones, see for example [40]. It is likely that they can be used to give the selection in more explicit form but this is not the aim of the paper. Below we provide a few examples of finding selections.

As an application of Theorems 3.8 and 4.3, we give a result concerning the convexity of weakly closed Tchebyshev sets in a reflexive Banach space whose norm is uniformly Gâteaux differentiable. The result was first established by V. Klee for X being a Banach space whose norm is both uniformly Gâteaux differentiable and uniformly rotund (the latter assumption implies the reflexivity of the space, see [18, Theorem 2, p. 27] and yields the Kadec–Klee property of the norm $\|\cdot\|$), see also [36, Corollary 4.2], and [1,49].

Theorem 4.5. *Let X be a Banach space whose norm is uniformly Gâteaux differentiable and $S \subset X$ be a closed set. Assume that either* (a) *or* (b) *or* (c) *below holds*:

(a) $\limsup_{\|h\|\to 0} \frac{d_S(x+h)-d_S(x)}{\|h\|} = 1$ *for all* $x \in X \setminus S$;
(b) $D_H d_S^2(x)$ *exists, and* $P_S(x) \neq \emptyset$ *for every* $x \in X$;
(c) *The space X is reflexive and the set S is weakly closed and Tchebyshev.*

Then the set S is convex and

$$E_{\frac{1}{2},X}\psi_S(x) = \{2J(x - P_S(x))\}.$$

Proof. Under (a), the convexity of S is justified by Theorem 4.3 and Proposition 4.2. If (b) holds, we deduce from Proposition 3.1 that the condition (a) is satisfied, so S is convex. Suppose now (c) holds. It follows from Theorem 3.8 applied to the function ψ_S (see Remark 3.1) that for every $x \in X$ we have

$$E_{\frac{1}{2},X}\psi_S(x) = \{2J(x - P_S(x))\}$$

and $D_H d^2(x)$ exists, so the convexity of S follows from (b). □

As a first consequence of (b) of the theorem above we have the following corollary.

Corollary 4.5. *Let X be a Banach space whose norm is uniformly Gâteaux differentiable and $S \subset X$ be a closed set such that $P_S(x) \neq \emptyset$ for all $x \in X$. Then S is convex if and only if anyone of properties* (a) *or* (b) *below holds*:

(a) *$\partial_C d_S(x)$ is a singleton for all $x \in X \setminus S$;*
(b) *d_S is Gâteaux differentiable at any point outside S.*

Proof. The implication (a) $\Rightarrow$ (b) is obvious and Theorem 4.5(b) ensures that the condition (b) of the corollary entails the convexity of the set S. On the other hand, Corollary 4.4(a) justifies that the convexity of S implies the condition (a) of the corollary. $\square$

The equivalence between the property (a) of the corollary and the convexity of S has been established in [19, Theorem 8] under the assumption that S is Tchebyshev and both norms $\|\cdot\|$ of the Banach space X and $\|\cdot\|_*$ of X^* are locally uniformly convex (which entails that both norms are also Fréchet differentiable off zero).

Another direct consequence of Theorem 4.5(b) and Corollary 4.3 is

Corollary 4.6. *Let H be a Hilbert space and $S \subset H$ be a nonempty closed subset such that*

$$H \setminus S \subset \bigcup_{x \in (\mathrm{Dom}\, P_S) \setminus S,\ z \in P_S(x)} [z, x[.$$

Then the set S is convex.

It is an interesting question whether the fulfillment of condition (a) in Corollary 4.4 implies the convexity of S. To the best of our knowledge this is an open problem in the case of Banach space whose norm $\|\cdot\|$ is uniformly Gâteaux differentiable. An answer to this question under some assumptions on X and X^* can be found in [19] for S being a Tchebyshev set. Below we give an answer to this question in Hilbert space setup.

Corollary 4.7. *Let $(H, \|\cdot\|)$ be a Hilbert space.*

(a) *If $S \subset X$ is a nonempty weakly closed set such that the set $\partial_C d_S(x)$ is singleton for every $x \in X \setminus S$, then S is convex.*
(b) *For any nonempty nonconvex weakly closed subset $S \subset X$, the squared distance function d_S^2 is a ∂_C-eds function but there is a point $u \in H$ at which it is not Clarke directionally subregular.*
(c) *If $S \subset X$ is a nonempty closed set such that the set $P_S(x) \neq \emptyset$ and $D_G d_S^2(x)$ exists for every $x \in X \setminus S$, then S is convex.*
(d) *If $S \subset X$ is a nonempty closed subset such that there exist a function f bounded from below by a quadratic function and a positive $\lambda \in]0, \frac{1}{2\alpha}[$ such that $e_\lambda f(\cdot) = -d_S^2(\cdot)$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$), then S is convex.*
(e) *If $S \subset X$ is a nonempty nonconvex closed subset, then $-d_S^2(\cdot)$ is a ∂_C-eds function which is not the Moreau envelope of a function f bounded from below by a quadratic function, i.e., there is no $\lambda \in]0, \frac{1}{2\alpha}[$ such that $e_\lambda f(\cdot) = -d_S^2(\cdot)$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$).*

Proof. (a) In order to prove the convexity of S we employ Theorem 4.5(c). It is enough to show that S is a Tchebyshev set. It follows from the reflexivity of the space and the weak closedness property of S that $P_S(x) \neq \emptyset$. Proposition 3.1 ensures that $P_S(x)$ is singleton, hence S is a Tchebyshev set, and by Theorem 4.5(c) S is convex.

The statement in (b) is consequence of Theorems 4.2 and 3.7 and of (a) above.

(c) The statement (c) follows directly from Theorem 4.5(b).

(d) Let f be a function bounded from below by a quadratic function and a positive $\lambda \in]0, \frac{1}{2\alpha}[$ be such that $e_\lambda f(\cdot) = -d_S^2(\cdot)$, where α is as in relation (3.5) (with the convention $\frac{1}{2\alpha} = +\infty$ for $\alpha = 0$). First observe that both functions h_1 and h_2 are convex, where

$$h_1(x) := \|x\|^2 - d_S^2(x) = \sup_{s \in S}\left\{\langle 2x, s\rangle - \|s\|^2\right\}$$

(is the Asplund function) and

$$h_2(x) := \frac{1}{2\lambda}\|x\|^2 + d_S^2(x) = \sup_{z \in H}\left\{\left\langle \frac{1}{\lambda}x, z\right\rangle - \frac{1}{2\lambda}\|z\|^2 - f(z)\right\}.$$

Take $x_1^* \in \partial h_1(x)$, $x_2^* \in -2\lambda\partial h_2(x)$ and observe that for every $y \in H$ we have

$$\|x+y\|^2 - \|x\|^2 - \langle x_1^*, y\rangle \geqslant d_S^2(x+y) - d_S^2(x) \geqslant \frac{1}{2\lambda}\left(\|x\|^2 - \|x+y\|^2 - \langle x_2^*, y\rangle\right),$$

which implies that $D_F d_S^2(x)$ exists (of course $\|D_F d_S^2(x)\| \leqslant 2d_S(x)$). In order to get the convexity of S, in view of (c), it is enough to show that $P_S(x) \neq \emptyset$ for every $x \in X \setminus S$. For this reason take sequences $\{t_i\}_{i\in\mathbb{N}}$ in $]0, 1[$ and $\{s_i\}_{i\in\mathbb{N}}$ in S such that $t_i \to 0$ and $d_S^2(x) + t_i^2 \geqslant \|x - s_i\|^2$ for every $i \in \mathbb{N}$ and $x - s_i \xrightarrow{w} v^* \in H$ (weakly). We have

$$\begin{aligned}
0 &\leqslant \lim_{i\to\infty} \frac{d_S^2(x + t_i(s_i - x)) - d_S^2(x) - \langle D_F d_S^2(x), t_i(s_i - x)\rangle}{t_i} \\
&\leqslant \lim_{i\to\infty} \frac{\|x + t_i(s_i - x) - s_i\|^2 - \|x - s_i\|^2 + t_i^2 - \langle D_F d_S^2(x), t_i(s_i - x)\rangle}{t_i} \\
&= -2d_S^2(x) - \langle D_F d_S^2(x), v^*\rangle \leqslant -2d_S^2(x) + \|D_F d_S^2(x)\|\,\|v^*\| \\
&\leqslant -2d_S^2(x) + 2d_S(x)\|v^*\| \leqslant -2d_S^2(x) + 2d_S^2(x) = 0,
\end{aligned}$$

so $\|v^*\| = d_S(x)$, which implies $\lim_{i\to\infty}\|x - s_i - v^*\| = 0$, so $x - v^* \in P_S(x)$. Thus, by (c) we get the convexity of S.

(e) Statements in (e) are simple consequences of (d) and Remark 3.3. □

In view of Theorem 4.5 items (a), (b) and (c) of the corollary above can be given in a more general setup, namely whenever the space is a reflexive Banach space whose norm is uniformly Gâteaux differentiable (the interested reader can obtain it by a repetition of reasonings from Theorem 4.5).

5. Representation of $\partial_L f(x)$ in terms of the Moreau envelope

In this section we turn our attention to the description of the Mordukhovich limiting subdifferential of a lower semicontinuous function f in terms of its Moreau envelope. A characterization of $\partial_L f(x)$ in terms of the limiting subdifferential of Moreau envelope was obtained in Theorem 5.5 in [33] in Asplund spaces. In the following theorem, we establish this connection in terms of Fréchet subdifferentials when the space is Asplund, and we also relate $\partial_L f(x)$ to Hadamard derivatives of Moreau envelope when the norm $\|\cdot\|$ of the Asplund space is in addition required to be uniformly Gâteaux differentiable. The latter property is enjoyed by super-reflexive Banach spaces, see [21, Corollary 1 and Corollary 3]. It should be noted that Theorem 5.1 was established in Theorem 3.10 [4] in the Hilbert space setting.

Before proving the theorem, let us give a lemma which will be used in the proof of the theorem. Although the lemma has its own interest, we do not find it (as stated below) in the literature.

Lemma 5.1. *Let $(X, \|\cdot\|)$ be a Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function. Then for any $\varepsilon \geqslant 0$*

$$\left\{\begin{array}{l}(x^*, -r^*) \in \partial_{F,\varepsilon}(\psi_{\operatorname{epi} f})(x,r) \\ \textit{and } r^* > \varepsilon\end{array}\right\} \quad \Rightarrow \quad r = f(x).$$

Proof. Take the sum norm $\|(x,r)\| := \|x\| + |r|$ on $X \times \mathbb{R}$ and fix $(x^*, -r^*) \in \partial_{F,\varepsilon}\psi_{\operatorname{epi} f}(x,r)$ with $r^* > \varepsilon$. By the definition of ε-Fréchet subdifferential, for any real $\varepsilon' > \varepsilon$ there exists some real $\delta > 0$ such that

$$\langle x^*, u - x\rangle - r^*(s - r) \leqslant \varepsilon' \big\|(u,s) - (x,r)\big\|$$

for all $(u,s) \in \operatorname{epi} f \cap B((x,r),\delta)$. Suppose that $r > f(x)$. We can then choose some real s such that $f(x) < s < r$ and $(x,s) \in B((x,r),\delta)$. Taking (x,s) in place of (u,s) in the above inequality yields

$$-r^*(s-r) \leqslant \varepsilon'(r-s) \quad \text{hence} \quad r^* \leqslant \varepsilon'.$$

Consequently $r^* \leqslant \varepsilon$, which contradicts the assumption $r^* > \varepsilon$. $\square$

Theorem 5.1. *Assume that $(X, \|\cdot\|)$ is an Asplund space and that f is lower semicontinuous and bounded from below by a negative quadratic function, see (3.6). Then for any x where f is finite one has*

$$\partial_L f(x) = \operatorname*{Lim\,sup}_{\substack{\lambda \downarrow 0 \\ (u, e_\lambda f(u)) \to (x, f(x))}} \partial_F e_\lambda f(u),$$

where Lim sup *means the weak* sequential outer (upper) limit superior of sets in X^*.*

If the norm $\|\cdot\|$ of the Asplund space X is in addition supposed to be uniformly Gâteaux differentiable, then one also has

$$\partial_L f(x) \subset \operatorname*{Lim\,sup}_{\substack{\lambda \downarrow 0 \\ (u, e_\lambda f(u)) \to (x, f(x))}} \big\{D_H e_\lambda f(u)\big\},$$

and the equality holds provided that f is weakly sequentially lower semicontinuous and $(X, \|\cdot\|)$ is Hilbert.

Proof. As in the proof of the above lemma, endow $X \times \mathbb{R}$ with the norm $\|(u,r)\| := \|u\| + |r|$, so the dual norm on $X^* \times \mathbb{R}$ is $\|(u^*,r^*)\|_* = \max\{\|u^*\|_*, |r^*|\}$. Fix any $x^* \in \partial_L f(x)$. By (2.3) and (2.5) there exists some real $t > 0$ such that

$$\frac{1}{t}\left(x^*, -1\right) \in \partial_L d\left(\operatorname{epi} f, \left(x, f(x)\right)\right).$$

Put $u^* = \frac{1}{t}x^*$ and $\alpha^* = -\frac{1}{t}$. Then by (2.6) there are sequences $\{(x_i,r_i)\}_{i\in\mathbb{N}}$ in $\operatorname{epi} f$ with $(x_i,r_i) \xrightarrow{\operatorname{epi} f} (x, f(x))$, $\{(x_i^*,\alpha_i^*)\}_{i\in\mathbb{N}}$ in $X^* \times \mathbb{R}$ with $(x_i^*,\alpha_i^*) \xrightarrow{w^*} (u^*,\alpha^*)$, $\{\varepsilon_i\}_{i\in\mathbb{N}}$ with $\varepsilon_i > 0$ and $\varepsilon_i \downarrow 0$, and $\{\delta_i\}_{i\in\mathbb{N}}$ with $0 < \delta_i < 1/4$ and $\delta_i \downarrow 0$ such that $\|(x_i^*,\alpha_i^*)\|_* \leqslant 1 + \varepsilon_i$ and

$$\begin{aligned} &-\left\langle x_i^*, u - x_i\right\rangle - \alpha_i^*(r - r_i) + \varepsilon_i \left\|(u,r) - (x_i,r_i)\right\| \geqslant 0, \\ &\quad \forall (u,r) \in \operatorname{epi} f \cap B\left[(x_i,r_i), \delta_i\right]. \end{aligned} \tag{5.1}$$

This ensures in particular $(x_i^*,\alpha_i^*) \in \partial_{F,\varepsilon_i}\psi_{\operatorname{epi} f}(x_i,r_i)$. Since $\alpha^* < 0$, by Lemma 5.1 for i sufficiently large we get $r_i = f(x_i)$. It follows from (3.8) that there are $0 < \lambda_i < 1$ such that $\lambda_i \downarrow 0$ and

$$C_i := \operatorname{epi} e_{\lambda_i} f \cap B\left[\left(x_i, f(x_i)\right), \frac{\delta_i}{2}\right] \subset \operatorname{epi} f + B\left(0, \delta_i^3\right), \tag{5.2}$$

recall $B[\cdot,\cdot]$ stands for the closed ball. We also observe that $C_i \neq \emptyset$ since $(x_i, f(x_i)) \in \operatorname{epi} e_{\lambda_i} f$ (keeping in mind the inclusion $\operatorname{epi} f \subset \operatorname{epi} e_\lambda f$ for any $\lambda > 0$). Combining relations (5.1) and (5.2) and putting $c_i = 1 + 2\varepsilon_i$, we get for i large enough

$$c_i\delta_i^3 - \left\langle x_i^*, u - x_i\right\rangle - \alpha_i^*\left(r - f(x_i)\right) + \varepsilon_i\left\|(u,r) - \left(x_i, f(x_i)\right)\right\| \geqslant 0, \quad \forall (u,r) \in C_i.$$

Consider the function

$$h_i(u,r) = -\left\langle x_i^*, u - x_i\right\rangle - \alpha_i^*\left(r - f(x_i)\right) + \varepsilon_i\left\|(u,r) - \left(x_i, f(x_i)\right)\right\|.$$

Then we have

$$h_i\left(x_i, f(x_i)\right) = 0 \leqslant h_i(u,r) + c_i\delta_i^3, \quad \forall (u,r) \in C_i$$

and $(x_i, f(x_i)) \in C_i$. By the Ekeland variational principle, there are $(x_i', r_i') \in C_i$ such that

$$\left\|\left(x_i', r_i'\right) - \left(x_i, f(x_i)\right)\right\| \leqslant \delta_i^2 \tag{5.3}$$

and

$$h_i\left(x_i', r_i'\right) \leqslant h_i(u,r) + c_i\delta_i\left\|(u,r) - \left(x_i', r_i'\right)\right\|, \quad \forall (u,r) \in C_i. \tag{5.4}$$

Thus for $(u,r) \in \operatorname{epi} e_{\lambda_i} f$ with $\|(u,r) - (x_i', r_i')\| < \delta_i^2$ we have

$$\begin{aligned} &-\left\langle x_i^*, x_i' - x_i\right\rangle - \alpha_i^*\left(r_i' - f(x_i)\right) + \varepsilon_i\left\|\left(x_i', r_i'\right) - \left(x_i, f(x_i)\right)\right\| \\ &\quad \leqslant -\left\langle x_i^*, u - x_i\right\rangle - \alpha_i^*\left(r - f(x_i)\right) + \varepsilon_i\left\|(u,r) - \left(x_i, f(x_i)\right)\right\| + c_i\delta_i\left\|(u,r) - \left(x_i', r_i'\right)\right\|, \end{aligned}$$

that is,

$$\begin{aligned}&\langle x_i^*, u - x_i'\rangle + \alpha_i^*(r - r_i')\\&\quad\leqslant \varepsilon_i \|(u,r) - (x_i, f(x_i))\| - \varepsilon_i \|(x_i', r_i') - (x_i, f(x_i))\| + c_i\delta_i \|(u,r) - (x_i', r_i')\|,\end{aligned}$$

hence

$$\langle x_i^*, u - x_i'\rangle + \alpha_i^*(r - r_i') \leqslant (\varepsilon_i + c_i\delta_i)\|(u,r) - (x_i', r_i')\|.$$

This implies that

$$(x_i^*, \alpha_i^*) \in \partial_{F,\varepsilon_i + c_i\delta_i} \psi_{\operatorname{epi} e_{\lambda_i} f}(x_i', r_i').$$

Since $\alpha^* < 0$, Lemma 5.1 yields $r_i' = e_{\lambda_i} f(x_i')$ for i large enough. Further, by the Asplund property of X and by (2.7) there are sequences $\{(u_i, s_i)\}_{i\in\mathbb{N}}$ in $X \times \mathbb{R}$ and $\{(u_i^*, s_i^*)\}_{i\in\mathbb{N}}$ in $X^* \times \mathbb{R}$ with $(u_i^*, s_i^*) \in N^F(\operatorname{epi} e_{\lambda_i} f, (u_i, s_i))$ such that

$$\|(u_i, s_i) - (x_i', e_{\lambda_i} f(x_i'))\| \leqslant \delta_i, \qquad \|(x_i^*, \alpha_i^*) - (u_i^*, s_i^*)\|_* \leqslant 2(\varepsilon_i + c_i\delta_i).$$

Since $s_i^* \to \alpha^*$ and $\alpha^* < 0$, by Lemma 5.1 we get $s_i = e_{\lambda_i} f(u_i)$ for i large enough and

$$\frac{-1}{s_i^*} u_i^* \in \partial_F e_{\lambda_i} f(u_i),$$

according to the first inclusion of (2.3). Since $u_i \to x$, $e_{\lambda_i} f(u_i) \to f(x)$, $s_i^* \to \alpha^*$, and

$$\frac{-1}{s_i^*} u_i^* \xrightarrow{w^*} \frac{-1}{\alpha^*} u^* = x^*,$$

we get

$$x^* \in \operatorname*{Lim\,sup}_{i\to+\infty} \partial_F e_{\lambda_i} f(u_i) \quad \text{and} \quad (u_i, e_{\lambda_i} f(u_i)) \to (x, f(x)),$$

which justifies the inclusion of the first member into the second of the desired equality.

Conversely, let $x^* \in \operatorname{Lim\,sup}_{\lambda\downarrow 0,\,(u, e_\lambda f(u))\to(x, f(x))} \partial_F e_\lambda f(u)$. Then there are sequences $\{\lambda_i\}_{i\in\mathbb{N}}$ with $\lambda_i > 0$ and $\lambda_i \downarrow 0$, $\{x_i\}_{i\in\mathbb{N}}$ in X with $x_i \to x$ and $e_{\lambda_i} f(x_i) \to f(x)$, and $\{x_i^*\}_{i\in\mathbb{N}}$ in X^* with $x_n^* \xrightarrow{w^*} x^*$, such that

$$x_i^* \in \partial_F e_{\lambda_i} f(x_i), \quad \text{for all } i \text{ sufficiently large.}$$

So there are sequences $\{\varepsilon_i\}_{i\in\mathbb{N}}$ with $\varepsilon_i > 0$ and $\varepsilon_i \downarrow 0$, and $\{\delta_i\}_{i\in\mathbb{N}}$ with $\delta_i > 0$ and $\delta_i \downarrow 0$ such that for i sufficiently large

$$e_{\lambda_i} f(x_i + h) - e_{\lambda_i} f(x_i) - \langle x_i^*, h\rangle + \varepsilon_i \|h\| \geqslant 0, \quad \forall h \in B[0, \delta_i].$$

Pick $v_i \in X$ such that

$$f(v_i) + \frac{1}{2\lambda_i}\|v_i - x_i\|^2 \leqslant e_{\lambda_i} f(x_i) + \delta_i^3.$$

The above inequality and relation (3.6) ensure that $v_i \to x$ and $\lim_{i\to+\infty} f(v_i) = f(x)$, and also

$$\delta_i^3 + e_{\lambda_i} f(x_i + h) - f(v_i) - \frac{1}{2\lambda_i}\|v_i - x_i\|^2 - \langle x_i^*, h\rangle + \varepsilon_i\|h\| \geqslant 0, \quad \forall h \in B[0, \delta_i].$$

Now taking into account that

$$e_{\lambda_i} f(x_i + h) \leqslant f(v_i + h) + \frac{1}{2\lambda_i}\|v_i - x_i\|^2$$

we get

$$\delta_i^3 + f(v_i + h) - f(v_i) - \langle x_i^*, h\rangle + \varepsilon_i\|h\| \geqslant 0, \quad \forall h \in B[0, \delta_i].$$

The Ekeland variational principle ensures the existence of $u_i \in X$ which satisfies

$$\|u_i\| \leqslant \delta_i^2 \quad \text{and} \quad f(v_i + u_i) - \langle x_i^*, u_i\rangle + \varepsilon_i\|u_i\| \leqslant f(v_i) \tag{5.5}$$

and for all $h \in B[0, \delta_i]$

$$f(v_i + h) - f(v_i + u_i) - \langle x_i^*, h - u_i\rangle + \varepsilon_i\|h\| - \varepsilon_i\|u_i\| + \delta_i\|h - u_i\| \geqslant 0,$$

which ensures

$$f(v_i + h) - f(v_i + u_i) - \langle x_i^*, h - u_i\rangle + \varepsilon_i\|h - u_i\| + \delta_i\|h - u_i\| \geqslant 0, \quad \forall h \in B[0, \delta_i].$$

By the first inequality in (5.5), for i large enough, u_i is an interior point of $B[0, \delta_i]$ thus

$$x_i^* \in \partial_{F,\varepsilon_i+\delta_i} f(v_i + \cdot)(u_i) = \partial_{F,\varepsilon_i+\delta_i} f(v_i + u_i),$$

so $x^* \in \partial_L f(x)$ since $\{(v_i + u_i, f(v_i + u_i))\}_{i\in\mathbb{N}}$ converges to $(x, f(x))$ according to the inequalities in (5.5), to the lower semicontinuity of f, and to the equality $\lim_{i\to\infty} f(v_i) = f(x)$. The first equality of the theorem is then proved.

Finally, if we suppose that the norm $\|\cdot\|$ of the Asplund space X is uniformly Gâteaux differentiable, the inclusion

$$\partial_L f(x) \subset \operatorname*{Lim\,sup}_{\substack{\lambda\downarrow 0 \\ (u, e_\lambda f(u))\to(x, f(x))}} \{D_H e_\lambda f(u)\}$$

as well the equality, under the additional assumptions in the theorem, follow from the fact that $\partial_F e_\lambda f(u) \subset \partial^- e_\lambda f(u)$ and from Theorem 3.5(d) and (f). □

The equality (2.11) then yields:

Corollary 5.1. *Assume in addition to the hypotheses of Theorem 5.1 that f is locally Lipschitz continuous near x_0. Then*

$$\partial_C f(x) = \overline{\mathrm{co}}^*\Big(\limsup_{\substack{\lambda \downarrow 0 \\ (u, e_\lambda f(u)) \to (x, f(x))}} \partial_F e_\lambda f(u) \Big).$$

Acknowledgment

We thank the referee for his comments on the presentation of the paper.

References

[1] E. Asplund, Čebyšev sets in Hilbert spaces, Trans. Amer. Math. Soc. 144 (1969) 235–240.
[2] H. Attouch, Convergence de fonctions convexes, des sous-différentiels et semi-groupes aassociés, C. R. Acad. Sci. Paris Sér. A–B 284 (1977) 539–542.
[3] H. Attouch, Variational Convergence for Functions and Operators, Pitman Advanced Publishing Program, Boston, 1984.
[4] M. Bacăk, J.M. Borwein, A. Eberhard, B.S. Mordukhovich, Infimal convolutions and Lipschitzian properties of subdifferentials for prox-regular functions in Hilbert spaces, J. Convex Anal. 17 (2010) 737–763.
[5] V.S. Balaganskii, L.P. Vlasov, The problem of convexity of Chebyshev sets, Russian Math. Surveys 51 (6) (1996) 1127–1190.
[6] F. Bernard, L. Thibault, Prox-regular functions in Hilbert spaces, J. Math. Anal. Appl. 303 (2005) 1–14.
[7] J.M. Borwein, Proximality and Chebyshev sets, Optim. Lett. (2007) 21–32.
[8] J.M. Borwein, S.P. Fitzpatrick, J.R. Giles, The differentiability of real functions on normed linear space using generalized subgradients, J. Math. Anal. Appl. 128 (1987) 512–534.
[9] M. Bougeard, Contributions à la théorie de Morse en dimension finie, Thesis, Université Paris IX Dauphine, 1978.
[10] H. Brézis, Propriétés régularisantes de certains semi-groupes non linéaires, Israel J. Math. 9 (1971) 513–534.
[11] H. Brézis, Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert, North-Holland, Amsterdam, 1973.
[12] R. Cibulka, M. Fabian, Attainment and (sub)differentiability of the infimal convolution of a function and the square of the norm, J. Math. Anal. Appl. 368 (2010) 538–550.
[13] F.H. Clarke, Optimization and Nonsmooth Analysis, A Wiley–Interscience Publication, 1983.
[14] R. Correa, A. Jofré, L. Thibault, Characterization of lower semicontinuous convex functions, Proc. Amer. Math. Soc. 116 (1992) 67–72.
[15] M. Degiovanni, A. Marino, M. Tosques, Evolution equations with lack of convexity, Nonlinear Anal. 9 (1985) 1401–1443.
[16] F.R. Deutsch, Best Approximation in Inner Product Spaces, Springer-Verlag, Berlin, 2001.
[17] R. Deville, G. Godefroy, V. Zizler, Smoothness and Renormings in Banach Spaces, Longman Scientific and Technical, 1993.
[18] J. Diestel, Geometry of Banach Spaces: Selected Topics, Lecture Notes in Math., vol. 485, Springer-Verlag, 1975.
[19] S. Dutta, Generalized subdifferential of the distance function, Proc. Amer. Math. Soc. 133 (2005) 2949–2955.
[20] I. Ekeland, J.-M. Lasry, On the number of periodic trajectories for a Hamilton flow, Ann. of Math. 112 (1980) 293–319.
[21] P. Enflo, Banach spaces which can be given an equivalent uniformly convex norm, Israel J. Math. 13 (1972) 281–288.
[22] M. Fabian, P. Hajek, V. Zizler, Uniform Eberlein compacta and uniformly Gâteaux smooth norm, Serdica Math. J. 23 (1997) 351–362.
[23] M. Fabian, G. Godefroy, V. Zizler, The structure of uniformly Gâteaux smooth Banach spaces, Israel J. Math. 124 (2001) 243–252.
[24] S. Fitzpatrick, Metric projections and the differentiability of distance functions, Bull. Aust. Math. Soc. 22 (1980) 291–312.

[25] S. Fitzpatrick, Differentiation of real-valued functions and continuity of metric projections, Proc. Amer. Math. Soc. 91 (1984) 544–548.
[26] J.R. Giles, Differentiability of distance functions and a proximal property inducing convexity, Proc. Amer. Math. Soc. 104 (1988) 458–464.
[27] J.R. Giles, A distance function property implying differentiability, Bull. Aust. Math. Soc. 39 (1989) 59–70.
[28] J.-B. Hiriart-Urruty, Potpourri of conjectures and open questions in nonlinear analysis and optimization, SIAM Rev. 49 (2007) 255–273.
[29] R.B. Holmes, Geometrical Functional Analysis and Applications, Springer-Verlag, New York, 1975.
[30] A.D. Ioffe, On the local surjection property, Nonlinear Anal. 11 (1987) 565–592.
[31] R.C. James, Super-reflexive Banach spaces, Canad. J. Math. 24 (1972) 896–904.
[32] M. Johanis, J. Rychtar, On uniformly Gâteaux smooth norms and normal structure, Proc. Amer. Math. Soc. 135 (2007) 1511–1514.
[33] A. Jourani, Limits superior of subdifferentials of uniformly convergent functions in Banach spaces, Positivity 3 (1999) 33–47.
[34] A. Jourani, L. Thibault, D. Zagrodny, $C^{1,\omega(\cdot)}$-regularity and Lipschitz-like properties of subdifferential, Proc. Lond. Math. Soc. 105 (2012) 189–223.
[35] A. Jourani, L. Thibault, D. Zagrodny, Moreau upper envelopes and their differential properties, in preparation.
[36] V. Klee, Convexity of Chebyshev sets, Math. Ann. 142 (1961) 292–304.
[37] K. Kuratowski, Une condition metrique pour la retraction des ensembles, C. R. Siences Soc. Sci. Lett. Varsovie 28 (1936) 156–158.
[38] S. Marcellin, L. Thibault, Evolution problems associated with primal lower nice functions, J. Convex Anal. 13 (2006) 385–421.
[39] E. Matoušková, How small are the sets where the metric projection fails to be continuous, Acta Univ. Carolin. Math. Phys. 33 (1992) 99–108.
[40] M. Mazade, Ensembles localement prox-réguliers et inéquations variationnelles, Doctor thesis, 2011.
[41] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation. I: Basic Theory, Grundlehren Math. Wiss. (Fundamental Principles of Mathematical Sciences), vol. 330, Springer-Verlag, Berlin, 2006, 584 pp.
[42] B.S. Mordukhovich, Y. Shao, Nonsmooth sequential analysis in Asplund spaces, Trans. Amer. Math. Soc. 348 (1996) 1235–1280.
[43] J.J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, C. R. Acad. Sci. Paris 255 (1962) 2897–2899.
[44] J.J. Moreau, Inf-convolution des fonctions numériques sur un espace vectoriel, C. R. Acad. Sci. Paris 256 (1963) 5047–5049.
[45] J.J. Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France 93 (1965) 273–299.
[46] J.-P. Penot, M. Bougeard, Approximation and decomposition properties of some classes locally d.c. functions, Math. Program. 41 (1988) 195–227.
[47] R.A. Poliquin, Integration of subdifferentials of nonconvex functions, Nonlinear Anal. 17 (1991) 385–398.
[48] R.A. Poliquin, R.T. Rockafellar, Prox-regular functions in variational analysis, Trans. Amer. Math. Soc. 348 (1996) 1805–1838.
[49] R.A. Poliquin, R.T. Rockafellar, L. Thibault, Local differentiability of distance functions, Trans. Amer. Math. Soc. 352 (2000) 5231–5249.
[50] B. Ricceri, A conjecture implying the existence of non-convex Chebyshev sets in infinite-dimensional Hilbert spaces, http://arxiv.org/abs/1101.2992v2, 2011.
[51] R.T. Rockafellar, R.J.-B. Wets, Variational Analysis, Grundlehren Math. Wiss., vol. 317, Springer, New York, 1998.
[52] J. Rychtář, Uniformly Gâteaux differentiable norms in spaces with unconditional basis, Serdica Math. J. 26 (2000) 353–358.
[53] L. Thibault, D. Zagrodny, Integration of subdifferentials of lower semicontinuous functions on Banach spaces, J. Math. Anal. Appl. 189 (1995) 33–58.
[54] L. Thibault, D. Zagrodny, Subdifferential determination of essentially directionally smooth functions in Banach space, SIAM J. Optim. 20 (2010) 2300–2326.
[55] L.P. Vlasov, Chebyshev sets and approximatively convex sets, Math. Notes 2 (1967) 600–605.
[56] L.P. Vlasov, Almost convexity and Chebyshev sets, Math. Notes 8 (1970) 776–779.
[57] L.P. Vlasov, Approximative properties of sets in normed linear spaces, Russian Math. Surveys 28 (6(174)) (1973) 3–66.
[58] X. Wang, On Chebyshev functions and Klee functions, J. Math. Anal. Appl. 368 (2010) 293–310.
[59] D. Zagrodny, The cancellation law for inf-convolution of convex functions, Studia Math. 110 (1994) 271–282.

[60] D. Zagrodny, Minimizers of the limit of Mosco converging functions, Arch. Math. 5 (2005) 440–445.
[61] D. Zagrodny, On the weak* convergence of subdifferentials of convex functions, J. Convex Anal. 12 (2005) 213–219.
[62] D. Zagrodny, A weak* approximation of subgradient of convex function, Control Cybernet. 36 (3) (2007) 793–802.
[63] L. Zajíček, Differentiability of the distance function and points of multivaluedness of the metric projection in Banach space, Czechoslovak Math. J. 33 (1983) 292–308.