Journal of Convex Analysis
Volume 30 (2023), No. 4, 1351–1378

Properties of the Fréchet Derivative of the Fenchel Conjugate of the Antidistance Function

Dariusz Zagrodny
Faculty of Mathematics and Computer Science, University of Lodz, Poland
dariusz.zagrodny@wmii.uni.lodz.pl

Dedicated to Roger J-B Wets on the occasion of his 85th birthday.

Received: February 27, 2022
Accepted: August 9, 2022

The Fréchet derivative of the Fenchel conjugate of the antidistance function is locally Lipschitz in the Hilbert space setting. It is shown that the Gâteaux differential of the operator, at points of the existence, is bounded self-adjoint operator (it is symmetric everywhere defined on the Hilbert space). Assuming Gâteaux or Fréchet differentiability of the operator results on the strong convergence of subgradients along some arcs are also provided.

Keywords: Antidistance function, subdifferentials, subgradients, Gâteaux derivative, Fréchet derivative.

2010 Mathematics Subject Classification: 49J52.

1. Introduction

Differentiability is a desired property of functions. The best known example of such a function, in the infinite dimensional Hilbert space setting, is the Moreau envelope, see [11] for several facts on the differentiability of the Moreau envelope. Except obvious transformations of norms of Hilbert spaces it is hard to point out functions which are differentiable on an open set. However, such functions should be identified. It turns out that the Fenchel conjugate of anitidistance function, say Δ_S^*, is $C^{1,1}$ locally on the open unit ball at the origin, see [23, Theorem 6.3] or Theorem 4.9 below, that is, the derivative of Δ_S^* is a locally Lipschitz operator, say T, see (18). Thus, due to results in [14], its derivative can be differentiable at points of a dense subset of this ball and hence an example of a function having the second order derivative at these points is provided, in the infinite Hilbert space setting. The second order theory of differentiability in infinite dimensional setting is not well recognized, in fact the state of the art is little. It seems that a progress in this direction can be made in a future, whenever a mass of examples will be gathered and explored in order to find common factors characterizing twice differentiable functions.

An important question concerning the second order derivative is its symmetry. However, even for Hessian matrices in the classical sense, symmetry can not be taken for granted, see [18, Chapter 13] for a detailed discussion on the subject in the finite dimensional setting. Herein this property is confirmed for Δ_S^*, see Fact 5.2 below, where A stands for the derivative of T at x^*. It is of interest that the directional

ISSN 0944-6532 / \$ 2.50

derivative of the antidistance function along directions from the image of A can be expressed by a simple formula involving A, see Fact 5.3. It is also revealed that the kernel of A contains the Minkowski difference of the subdifferential of Δ_S at $T(x^*)$, that is $\partial\Delta_S(T(x^*)) - \partial\Delta_S(T(x^*)) \subset \ker A$, see Fact 5.3. Thus if the kernel is finite dimensional, the strong convergence of subgradients along some curves can be preserved, see for example Proposition 6.2 or Corollary 6.4; see also [7] and [25, 26] for questions concerning the convergence of subgradients.

The author would like to thank to Professor Lionel Thibault for several remarks improving the presentation of this work.

2. Preliminaries

In this section some basic notions and their properties are gathered.

In the sequel $\mathbb{H}$ stands for a real Hilbert space (with a real inner product $\langle\cdot,\cdot\rangle$), see for example [24]. If $Y \subset \mathbb{H}$ is a linear closed subspace of $\mathbb{H}$ then for every real $r > 0$ and every $x \in Y$ we denote by $\mathbb{B}_Y(x,r)$ (resp. $\mathbb{B}_Y[x,r]$) the open (resp. closed) ball centered at x and of radius r in Y, the sphere is denoted by $\mathbb{S}_Y[x,r] := \{y \in \mathbb{H} \mid \|y-x\| = r\}$; in the case $Y = \mathbb{H}$ the notations $\mathbb{B}(x,r)$, $\mathbb{B}[x,r]$ and $\mathbb{S}[x,r]$ are used respectively. For a subset $D \subset \mathbb{H}$, "$\operatorname{int} D$" stands for the interior and "$\operatorname{cl} D$" for the topological closure of D, a point x is an interior point of D if there exists an open ball centered at x which is completely contained in D. The convex hull of a set D is denoted by $\operatorname{conv} D$, see [4] for example, and $\operatorname{span} D$ stands for the smallest linear subspace containing D, see [1, p. 1].

In the sequel the following two classes of functions are used:

$$\mathscr{F} := \{f\colon [0,\infty[\to [0,\infty[\cup\{\infty\} \mid f(0)=0, \forall t\in]0,\infty[,\ f(t)>0,\ \lim_{t\to 0} f(t)=0\} \quad (1)$$

and

$$\mathscr{G} := \{g\colon [0,\infty[\to [0,\infty[\cup\{\infty\} \mid \forall t\in]0,\infty[,\ g(t)>0,\ \limsup_{t\to 0} g(t)<\infty\}. \quad (2)$$

For subsets $C, D \subset \mathbb{H}$, $C \subset D$, normed space $(Z, |||\cdot|||)$ and some $M \geq 0$ a given operator $F : D \longrightarrow Z$ is said to be M-Lipschitz on C if $|||F(a)-F(b)||| \leq M\|a-b\|$, whenever $a, b \in C$. F is said to be locally Lipschitz on $B \subset D$ if for every $b \in B$ there is an open subset $U \subset \mathbb{H}$ such that $b \in U$ and F is M-Lipschitz on $B \cap U$ for some $M \geq 0$. If $d \in \operatorname{int} D$ and $v \in \mathbb{H}$ is given then the directional derivative of F at d along v is defined as

$$F'(d;v) := \lim_{t\longrightarrow 0} \frac{F(d+tv)-F(d)}{t},$$

whenever the limit exists. We recall that F is *Gâteaux differentiable* at $d \in \operatorname{int} D$ provided there exists a continuous linear operator $A : \mathbb{H} \longrightarrow Z$ ($D_G F(d) := A$) such that for all $h \in \mathbb{H}$ the following equality

$$\lim_{t\downarrow 0} t^{-1}(F(d+th)-F(d)) = \langle A, h\rangle \quad (3)$$

holds true. If, for a fixed d, the convergence in (3) is uniform with respect to h on every bounded subset, then we say that F is *Fréchet differentiable* in d and the derivative is denoted by $D_F F(d)$.

Let $f : \mathbb{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous convex function which is finite at $x \in \mathbb{H}$. The *subdifferential* of f at $x \in \mathbb{H}$ is defined by

$$\partial f(x) := \{x^* \in \mathbb{H} \mid \forall h \in \mathbb{H}, \langle x^*, h\rangle \leq f(x+h) - f(x)\},$$

see [1, Definition 16.1] for example. The nonemptiness of subdifferentials at a dense subset of the domain of f is guaranteed by the celebrated Brøndsted-Rockafellar Theorem, see [17] for a detailed discussion. Below, following [17, Theorem 3.17], a version of it is recalled.

Theorem 2.1. *Suppose that $f : \mathbb{H} \longrightarrow \mathbb{R}\cup\{+\infty\}$ is a lower semicontinuous convex function. Then given any point $x_0 \in \mathbb{H}$, with $f(x_0) < \infty$, $\varepsilon > 0$, $\lambda > 0$ and $x_0^* \in \mathbb{H}$ such that*

$$\forall h \in \mathbb{H}, \langle x_0^*, h\rangle \leq \varepsilon\lambda + f(x_0 + h) - f(x_0),$$

there exist $x \in \mathbb{H}$, $x^* \in \mathbb{H}$ such that

$$x^* \in \partial f(x), \|x - x_0\| \leq \varepsilon \text{ and } \|x^* - x_0^*\| \leq \lambda.$$

Below a result ensuring the weak* convergence of subgradients along some directions is recalled, see [25].

The weak convergence is denoted by $\xrightarrow{weak}$, see for example [1, 2.4 Strong and Weak Topologies]. Let us recall that any bounded sequence of elements of a Hilbert space has a weak sequential cluster point, see for example [1, Fact 2.27 (Banach-Alaoglu), Fact 2.30]. Information on the weak topology can be found in [21, 24] too.

Lemma 2.2. *Let $(X, \|\cdot\|)$ be a reflexive Banach space and $f : X \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a convex lower semicontinuous function, $x \in \operatorname{dom} f$ and $M \geq 0$, $w_0 \in X$, $x^* \in X^*$ be given. If $f'(x; w_0)$ is a finite real number (the directional derivative at x along w_0 is finite) such that*

$$\forall h \in \mathbb{S}_X[0,1],\ \lim_{t\downarrow 0} \frac{f'(x; w_0 + th) - f'(x; w_0) - \langle x^*, th\rangle}{t} = 0, \tag{4}$$

then for all sequences $\{t_i\}_{i\in\mathbb{N}}$ such that $t_i > 0$, $\mathbb{B}_{X^*}[0, M] \cap \partial f(x + t_i w_0) \neq \emptyset$ for all $i \in \mathbb{N}$, and $t_i \downarrow 0$ we have

$$w - \lim_{i\longrightarrow\infty} \mathbb{B}_{X^*}[0, M] \cap \partial f(x + t_i w_0) = \{x^*\}, \tag{5}$$

and consequently
$$w - \lim_{t\downarrow 0} \mathbb{B}_{X^*}[0, M] \cap \partial f(x + t w_0) = \{x^*\}, \tag{6}$$

whenever $\mathbb{B}_{X^*}[0, M] \cap \partial f(x + t w_0) \neq \emptyset$ for all $t \in]0, \delta[$ for some $\delta > 0$, where

$$w - \lim_{i\longrightarrow\infty} \mathbb{B}_{X^*}[0, M] \cap \partial f(x + t_i w_0) := \tag{7}$$
$$\left\{y^* \in X^* : x_i^* \xrightarrow{weak} y^*,\ \text{with}\ \ x_i^* \in \mathbb{B}_{X^*}[0, M] \cap \partial f(x + t_i w_0)\, \forall\ i \in \mathbb{N}\right\}.$$

Moreover
$$f'(x; w_0) = \langle x^*, w_0\rangle. \tag{8}$$

The *Fenchel conjugate* of f, denoted by f^*, is defined by

$$\forall x^* \in \mathbb{H}, f^*(x^*) := \sup_{z\in\mathbb{H}} \langle x^*, z\rangle - f(z),$$

see [1, Chapters 14–16] for further information on this notion. It is well-known that $f^* : \mathbb{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a convex lower semicontinuous function, whenever $f(u) \in \mathbb{R}$ for some $u \in \mathbb{H}$. Moreover, if $f(x) \in \mathbb{R}$ then

$$x^* \in \partial f(x) \Longleftrightarrow f(x) + f^*(x^*) = \langle x^*, x \rangle, \tag{9}$$

see [1, Theorem 16.23] for details.

A subset $G \subset \mathbb{H} \times \mathbb{H}$ is said to be *monotone* provided that

$$\langle x^* - y^*, x - y \rangle \geq 0, \text{ whenever } (x, x^*), (y, y^*) \in G.$$

The subset G is said to be *maximally monotone* if G is monotone and G has no proper extension (subset G is maximal in the family of monotone subsets of $\mathbb{H} \times \mathbb{H}$, ordered by inclusion), see [1, Chapter 20] for more on monotone operators in the Hilbert space setting. It is known that the graph of the subdifferential of a proper lower semicontinuous function $f : \mathbb{H} \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a maximal monotone subset of $\mathbb{H} \times \mathbb{H}$, that is the set

$$G(\partial f) := \{(x, x^*) \in \mathbb{H} \times \mathbb{H} \mid x^* \in \partial f(x)\},$$

is maximal monotone, whenever $f(u) \in \mathbb{R}$ for some $u \in \mathbb{H}$, see for example [4, Theorem 3.5.3]. It is a simple consequence of the Perfect Square Criterion for Maximality, see [22, Theorem 10.3], that for any $(w, w^*) \in \mathbb{H} \times \mathbb{H}$ there is some $(v, v^*) \in G(\partial f)$ such that

$$w + w^* = v + v^*. \tag{10}$$

Assuming that $S \subset \mathbb{H}$ is bounded the *farthest distance function* from the set S (also called antidistance function) is defined by

$$\Delta_S(x) := \sup_{y \in S} \|x - y\|, \tag{11}$$

and the set $Q_S(x)$ of farthest points (the attainment set) from S to x is given as

$$Q_S(x) := \{s \in S : \|x - s\| = \Delta_S(x)\}. \tag{12}$$

In several important cases the set of farthest points of S and of the closure of the convex hull of S, $\operatorname{cl}\operatorname{conv} S$, coincide, that is $Q_S(x) = Q_{\operatorname{cl}\operatorname{conv} S}(x)$, see [12, Theorem 2] for example. This allows us to assume, without of loss generality, that S is closed and convex, whenever antidistance function is under consideration. We put also

$$\operatorname{diam} S := \sup_{x,y \in S} \|x - y\|.$$

The *metric projection mapping* on $B \subset \mathbb{H}$ is defined by

$$P_B(x) := \{s \in \operatorname{cl} B \mid d_B(x) = \|x - s\|\}; \tag{13}$$

the set on the right side of the equality can be empty, where d_B stands for the distance function to B, that is $d_B(x) := \inf_{b \in B} \|x - b\|$.

Let $A : \mathbb{H} \longrightarrow \mathbb{H}$ be a bounded linear operator, see for example [24, Definiton 1, p. 43] see also [8, Chapter 8]. The image of A is denoted by $A(\mathbb{H})$ and its kernel by $\ker A$, that is,

$$A(\mathbb{H}) := \{Ay \mid y \in \mathbb{H}\} \text{ and } \ker A := \{x \in \mathbb{H} \mid Ax = 0\}.$$

The operator A is said to be *self-adjoint* if

$$\forall x, y \in \mathbb{H},\ \langle x, Ay\rangle = \langle Ax, y\rangle,$$

see for example [24, Definition 2, p. 197], see also [8, Chapter 8, Section 8.3]. For any linear subspace $X \subset \mathbb{H}$ the *orthogonal complement* of X is the set

$$X^{\perp} := \{y \in \mathbb{H} \mid \forall x \in X, \langle x, y\rangle = 0\},$$

see [24, Theorem 1, p. 82] see also [8, Chapter 7, Section 7.5]. It follows from [24, Theorem 1, p. 82] that $\mathbb{H} = X + X^{\perp}$, whenever X is a closed linear subspace.

Fact 2.3. *Let $A : \mathbb{H} \longrightarrow \mathbb{H}$ be a bounded linear self-adjoint operator, then*

$$A(\mathbb{H})^{\perp} = \ker A.$$

Moreover, if A has closed image, that is, $\operatorname{cl} A(\mathbb{H}) = A(\mathbb{H})$, *then*

$$\{x \in \mathbb{H} \mid Ax = 0\}^{\perp} = A(\mathbb{H})$$

and for any sequence $\{p_i\}_{i\in\mathbb{N}}$ of elements of $\mathbb{H}$ we have

$$\lim_{i\to\infty} \|Ap_i\| = 0 \Longrightarrow \lim_{i\to\infty} \|P_{A(\mathbb{H})}(p_i)\| = 0.$$

Proof. Proofs of the equalities $A(\mathbb{H})^{\perp} = \ker A$ and $\{x \in \mathbb{H} \mid Ax = 0\}^{\perp} = A(\mathbb{H})$ (which are commonly known) can be done by a simple algebra, so they are omitted here. Let us fix any sequence $\{p_i\}_{i\in\mathbb{N}}$ elements of $\mathbb{H}$ such that $\lim_{i\to\infty} \|Ap_i\| = 0$ and assume that $A(\mathbb{H})$ is a closed linear subspace of $\mathbb{H}$. It follows from the Open Mapping Theorem, see [24, Theorem 1, p. 75], that there is $\delta > 0$ for which

$$\{y \in A(H) \mid \|y\| < \delta\} \subset \{Ax \mid \|x\| < 1\}.$$

Hence

$$\begin{aligned} 0 &= \lim_{i\to\infty} \|Ap_i\| = \lim_{i\to\infty} \left(\sup_{x\in\mathbb{B}(0,1)} \langle x, Ap_i\rangle \right) = \lim_{i\to\infty} \left(\sup_{x\in\mathbb{B}(0,1)} \langle Ax, p_i\rangle \right) \\ &= \lim_{i\to\infty} \left(\sup_{x\in\mathbb{B}(0,1)} \langle Ax, P_{A(\mathbb{H})}(p_i)\rangle \right) \geq \lim_{i\to\infty} \left(\sup_{v\in\mathbb{B}(0,\delta)} \langle v, P_{A(\mathbb{H})}(p_i)\rangle \right) \\ &= \delta \lim_{i\to\infty} \|P_{A(\mathbb{H})}(p_i)\| \geq 0. \end{aligned}$$

□

3. The subdifferential of the antidistance function

Let $S \subset \mathbb{H}$ be a bounded subset. For any $\varepsilon > 0$ and $x \in \mathbb{H}$ put

$$Q_{S,\varepsilon}(x) := \{s \in S \mid \|x - s\| + \varepsilon \geq \Delta_S(x)\}. \tag{14}$$

In this section a characterization of the subdifferential of the antidistance function Δ_S, defined in (11), by the set $Q_{S,\varepsilon}(x)$ is provided. For the distance function a characterization of its subdifferential was presented in [11, Theorem 3.2], in a more general case. Formulae given in (15) and that from [11, Theorem 3.2] rely on the same idea in order to perform subgradients as a limit of convex combinations of elements related to almost farthest points. A formulae for the subdifferential of antidistance function involving the farthest points of elements from a neighbourhoods of x can be found for example in [23, Proposition 4.3].

Proposition 3.1. *Let $S \subset \mathbb{H}$ be a bounded subset. Then for each $x \in \mathbb{H}$ such that $\Delta_S(x) > 0$ we have:*

(i)
$$\Delta_S(x)\partial\Delta_S(x) = \bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv}(x - Q_{S,\varepsilon}(x))\,; \tag{15}$$

(ii) *for every $h \in \mathbb{H}$ there is a sequence $\{s_i\}_{i\in\mathbb{N}}$ of elements of S such that*
$$\lim_{i\to\infty} \|x - s_i\| = \Delta_S(x) \quad \textit{and} \quad \Delta'_S(x;h) \le \liminf_{i\to\infty} \frac{\langle x - s_i, h\rangle}{\Delta_S(x)};$$

(iii) *if $\partial\Delta_S(x) \subset \mathbb{S}[0,1]$ and S is additionally closed and convex, then both sets $\partial\Delta_S(x)$ and $Q_S(x)$ are singletons, and $D_F\Delta_S(x)$ exists.*

Proof. Proof of the inclusion "$\supset$" in (i): First of all let us observe that by (14) the sets $\operatorname{cl}\operatorname{conv}(x - Q_{S,\varepsilon}(x))$ are weakly compact, see [21, Weak Topologies, The Banach-Alaoglu Theorem] for example, hence the intersection on the right side of the equality in (15) is nonempty. Suppose that $x^* \in \bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv}(x - Q_{S,\varepsilon}(x))$. For all $\varepsilon > 0$ let us choose $k(\varepsilon) \in \mathbb{N}$, $\lambda_1 > 0, \dots, \lambda_{k(\varepsilon)} > 0$, $a_1 \in S\backslash\{x\}, \dots, a_{k(\varepsilon)} \in S\backslash\{x\}$ such that
$$\sum_{i=1}^{k(\varepsilon)} \lambda_i = 1,\ \Delta_S(x) \le \sum_{i=1}^{k(\varepsilon)} \lambda_i\|x - a_i\| + \varepsilon,\ \|x^* - \sum_{i=1}^{k(\varepsilon)} \lambda_i(x - a_i)\| \le \varepsilon.$$

Thus for all $h \in H$ and $\varepsilon \in]0, \Delta_S(x)[$ we get
$$\begin{aligned}
\Delta_S(x+h) - \Delta_S(x) &\ge -\varepsilon + \sum_{i=1}^{k(\varepsilon)} \lambda_i\Big(\|x+h-a_i\| - \|x - a_i\|\Big)\\
&\ge -\varepsilon + \langle\sum_{i=1}^{k(\varepsilon)} \frac{\lambda_i}{\|x-a_i\|}(x-a_i), h\rangle \ge -\varepsilon + \langle\sum_{i=1}^{k(\varepsilon)} \frac{\lambda_i}{\Delta_S(x)}(x-a_i), h\rangle\\
&- \|h\|\left(1 - \sum_{i=1}^{k(\varepsilon)} \frac{\lambda_i\|x-a_i\|}{\Delta_S(x)}\right) \ge -\varepsilon\left(1 + 2\frac{\|h\|}{\Delta_S(x)}\right) + \frac{\langle x^*, h\rangle}{\Delta_S(x)}.
\end{aligned}$$

Now $x^* \in \Delta_S(x)\partial\Delta_S(x)$ follows after passing to the limit as $\varepsilon \downarrow 0$.

(ii) Let us fix some $h \in \mathbb{H}$. For all $i \in \mathbb{N}$ let us choose $s_i \in S \setminus \{x\}$ such that
$$\forall i \in \mathbb{N}, \|x + i^{-1}h - s_i\| + i^{-2} \ge \Delta_S(x + i^{-1}h)$$

and observe that
$$\begin{aligned}
\Delta'_S(x,h) &= \lim_{i\to\infty} i\Big(\Delta_S(x+i^{-1}h) - \Delta_S(x)\Big) \le \lim_{i\to\infty} i\Big(\|x+i^{-1}h-s_i\| + i^{-2} - \|x-s_i\|\Big)\\
&\le \liminf_{i\to\infty} \frac{\langle x-s_i, h\rangle}{\|x-s_i\|} = \liminf_{i\to\infty} \frac{\langle x-s_i, h\rangle}{\Delta_S(x)}.
\end{aligned}$$

Proof of the inclusion "$\subset$" in (i): We fix some $y^* \notin \bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv}(x - Q_{S,\varepsilon}(x))$. There is $\varepsilon > 0$ such that $y^* \notin \operatorname{cl}\operatorname{conv}(x - Q_{S,\varepsilon}(x))$, so there are $\delta > 0$ and $w \in \mathbb{H}$ for which
$$\forall s \in Q_{S,\varepsilon}(x), \langle y^*, w\rangle > \delta + \langle x - s, w\rangle.$$

For all $i \in \mathbb{N}$ let us choose $s_i \in S \setminus \{x\}$ such that

$$\lim_{i\to\infty} \|x - s_i\| = \Delta_S(x), \Delta'_S(x, w) \le \liminf_{i\to\infty} \frac{\langle x - s_i, w\rangle}{\Delta_S(x)}$$

and observe that

$$\Delta'_S(x; w) \le \liminf_{i\to\infty} \frac{\langle x - s_i, w\rangle}{\|x - s_i\|} \le \liminf_{i\to\infty} \frac{\langle y^*, w\rangle - \delta}{\|x - s_i\|} = \frac{\langle y^*, w\rangle - \delta}{\Delta_S(x)}.$$

Hence $y^* \notin \Delta_S(x)\partial\Delta_S(x)$ and finally $\Delta_S(x)\partial\Delta_S(x) = \bigcap_{\varepsilon>0} \mathrm{cl\,conv}\,(x - Q_{S,\varepsilon}(x))$.

(iii) if $\partial\Delta_S(x) \subset \mathbb{S}[0, 1]$ and S is bounded, closed and convex, then it follows from (15) that

$$\bigcap_{\varepsilon>0} \mathrm{cl\,conv}\,(x - Q_{S,\varepsilon}(x))$$

is a singleton, say $\{p\} = \bigcap_{\varepsilon>0} \mathrm{cl\,conv}\,(x - Q_{S,\varepsilon}(x))$, otherwise by the convexity of $\partial\Delta_S(x)$ we get $\partial\Delta_S(x) \cap \mathbb{B}(0, 1) \neq \emptyset$, a contradiction. Obviously $x - p \in S$ and $\Delta_S(x) = \|p\|$, and $\partial\Delta_S(x)$ must be a singleton, since $\bigcap_{\varepsilon>0} \mathrm{cl\,conv}\,(x - Q_{S,\varepsilon}(x))$ is a singleton. Let us also notice that if $s_i \in S$ for all $i \in \mathbb{N}$ and $\lim_{i\to\infty} \|x - s_i\| = \Delta_S(x)$, then $\lim_{i\to\infty} \|x - s_i - p\| = 0$, since $\bigcap_{\varepsilon>0} \mathrm{cl\,conv}\,(x - Q_{S,\varepsilon}(x))$ is a singleton. Thus taking any $h_i \in \mathbb{H} \setminus \{0\}$, $s_i \in S$ such that $\lim_{i\to\infty} \|h_i\| = 0$ and

$$0 < \Delta_S(x + h_i) - \|h_i\|^2 \le \|x + h_i - s_i\| \quad \text{for all} \quad i \in \mathbb{N},$$

we obtain for all $i \in \mathbb{N}$,

$$\begin{aligned} 0 &\le \Delta_S(x + h_i) - \Delta_S(x) - \frac{\langle p, h_i\rangle}{\Delta_S(x)} \le \|x + h_i - s_i\| + \|h_i\|^2 - \|x - s_i\| - \frac{\langle p, h_i\rangle}{\Delta_S(x)} \\ &\le \frac{\langle x + h_i - s_i, h_i\rangle}{\|x + h_i - s_i\|} - \frac{\langle p, h_i\rangle}{\Delta_S(x)} + \|h_i\|^2, \end{aligned}$$

which implies
$$\lim_{i\to\infty} \frac{\left| \Delta_S(x + h_i) - \Delta_S(x) - \frac{\langle p,h_i\rangle}{\Delta_S(x)} \right|}{\|h_i\|} = 0,$$

so that the Fréchet derivative $D_F\Delta_S(x)$ exists. □

The property presented in (iii) can be also inferred from [6, Corollary 3.6.(b)], where it was presented under assumptions of differentiability of norms of the space and its dual.

Corollary 3.2. *Let $S \subset \mathbb{H}$ be a bounded subset. Then for each $x \in \mathbb{H}$ such that $\Delta_S(x) > 0$ we have:*

(i) $\forall y \in \mathbb{H} \setminus \{x\}, \Delta^2_S(y) - \Delta^2_S(x) \ge \|y - x\|^2 + 2\Delta_S(x)\Delta'_S(x; y - x)$;

(ii) *let $\beta \ge 0$, $w \in \mathbb{H}$ and let the sequences $\{\varepsilon_i\}_{i\in\mathbb{N}}$, $\{z_i\}_{i\in\mathbb{N}}$ be such that $\varepsilon_i \searrow 0$, $\|w - z_i\| + \varepsilon_i \ge \Delta_S(w)$ and*

$$z_i \in S \cap \Big(\mathrm{cl\,conv}\, Q_{S,\varepsilon_i}(x) + \{v \in \mathbb{H} \mid \langle v, w - x\rangle \ge -\beta\}\Big)$$

for all $i \in \mathbb{N}$, then

$$\Delta^2_S(w) - \Delta^2_S(x) \le \beta + \|w - x\|^2 + 2\Delta_S(x)\Delta'_S(x; w - x);$$

in particular, if $S \subset \bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv} Q_{S,\varepsilon}(x)$, *then*

$$\forall y \in \mathbb{H}, \Delta_S^2(y) - \Delta_S^2(x) = \|y-x\|^2 + 2\Delta_S(x)\Delta_S'(x; y-x);$$

(iii) *if* $\partial\Delta_S(x) \cap \mathbb{B}(0,1) \neq \emptyset$, *then* $\partial\Delta_S(x) \cap \mathbb{B}(0,1) \cap \bigcup_{y\in\mathbb{H}\setminus\{x\}} \partial\Delta_S(y) = \emptyset$;

(iv) *if we have* $x^* \in \partial\Delta_S(x) \cap \mathbb{B}(0,1)$, $y^* \in \partial\Delta_S(y)$ *satisfying* $\|x^*\| + \varepsilon \leq 1$ *and* $\Delta_S(x) + \Delta_S(y) \leq \eta\varepsilon$ *for some* $\varepsilon > 0$, $\eta > 0$, *then*

$$\|y-x\|^2 \leq \eta\langle y^* - x^*, y-x\rangle;$$

(v) *the set* $\operatorname{argmin} \Delta_S$ *is a singleton, where*

$$u \in \operatorname{argmin} \Delta_S \implies 0 \in \partial\Delta_S(u).$$

Proof. (i) Let us take some $y \in \mathbb{H}$ and fix a sequence $\{s_i\}_{i\in\mathbb{N}}$ of elements of S such that

$$\lim_{i\to\infty} \|x-s_i\| = \Delta_S(x) \text{ and } \Delta_S'(x; y-x) \leq \liminf_{i\to\infty} \frac{\langle x-s_i, y-x\rangle}{\Delta_S(x)},$$

see Proposition 3.1(ii). Thus

$$\begin{aligned}\Delta_S^2(y) &\geq \liminf_{i\to\infty} \|y-s_i\|^2 = \|y-x\|^2 + \liminf_{i\to\infty}\Big(2\langle x-s_i, y-x\rangle + \|x-s_i\|^2\Big)\\ &\geq \|y-x\|^2 + 2\Delta_S(x)\Delta_S'(x; y-x) + \Delta_S^2(x).\end{aligned}$$

(ii) Let $\beta \geq 0$, $w \in \mathbb{H}$ and let the sequences $\{\varepsilon_i\}_{i\in\mathbb{N}}$, $\{z_i\}_{i\in\mathbb{N}}$ be such that $\varepsilon_i \searrow 0$, $\|w-z_i\| + \varepsilon_i \geq \Delta_S(w)$ and $z_i \in S \cap \big(\operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x) + \{v \in \mathbb{H} \mid \langle v, w-x\rangle \geq -\beta\}\big)$. Thus we have for all $i \in \mathbb{N}$: $\|x-p_i\| \leq \Delta_S(x)$, $z_i = p_i + v_i$, $p_i \in \operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x)$ and $v_i \in \{v \in \mathbb{H} \mid \langle v, w-x\rangle \geq -\beta\}$, and

$$\begin{aligned}\Delta_S^2(w) &= \lim_{i\to\infty} \|(w-x)+(x-z_i)\|^2 = \lim_{i\to\infty}\big(\|w-x\|^2 + 2\langle w-x, x-z_i\rangle + \|x-z_i\|^2\big)\\ &\leq \|w-x\|^2 + \Delta_S^2(x) + 2\lim_{i\to\infty}\langle w-x, x-p_i-v_i\rangle\\ &\leq \|w-x\|^2 + \Delta_S^2(x) + 2\lim_{i\to\infty}\sup_{s\in\operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x)}\langle w-x, x-s\rangle + \beta.\end{aligned}$$

Taking any sequence $\{s_i\}_{i\in\mathbb{N}}$ such that $s_i \xrightarrow{weak} s^*$ and

$$s_i \in \operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x), \quad \sup_{s\in\operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x)}\langle w-x, x-s\rangle \leq i^{-1} + 2\langle w-x, x-s_i\rangle$$

for all $i \in \mathbb{N}$, we get $v^* := \frac{1}{\Delta_S(x)}(x-s^*) \in \partial\Delta_S(x)$, see (15), and

$$\Delta_S^2(w) \leq \beta + \|w-x\|^2 + \Delta_S^2(x) + 2\Delta_S(x)\langle w-x, v^*\rangle.$$

Thus the inequality $\langle w-x, v^*\rangle \leq \Delta_S'(x; w-x)$ yields the following inequality

$$\Delta_S^2(w) - \Delta_S^2(x) \leq \beta + \|w-x\|^2 + 2\Delta_S(x)\Delta_S'(x; w-x).$$

Moreover, the inclusion $S \subset \bigcap_{\varepsilon>0} \operatorname{cl}\operatorname{conv} Q_{S,\varepsilon}(x)$ ensures that for every $y \in \mathbb{H}$ and any sequences $\{z_i\}_{i\in\mathbb{N}}$, $\{\varepsilon_i\}_{i\in\mathbb{N}}$ such that $\varepsilon_i \searrow 0$, $\|y-z_i\| + \varepsilon_i \geq \Delta_S(y)$ and $z_i \in S$ for all $i \in \mathbb{N}$ we have $\forall i \in \mathbb{N}, z_i \in \operatorname{cl}\operatorname{conv} Q_{S,\varepsilon_i}(x)$.

Thus for all $y \in \mathbb{H}$, $\Delta_S^2(y) - \Delta_S^2(x) \leq \|y-x\|^2 + 2\Delta_S(x)\Delta_S'(x; y-x)$, which, by (i) implies

$$\forall y \in \mathbb{H}, \Delta_S^2(y) - \Delta_S^2(x) = \|y-x\|^2 + 2\Delta_S(x)\Delta_S'(x; y-x).$$

(iii) Let $x^* \in \partial\Delta_S(x) \cap \mathbb{B}(0,1)$. Suppose that $x^* \in \partial\Delta_S(y)$ and $x \neq y$, then $\Delta_S(y) = \Delta_S(x) + \langle x^*, y-x\rangle$. It follows from (i) that

$$\begin{aligned} \|y-x\|^2 + 2\Delta_S(x)\langle x^*, y-x\rangle &> (\langle x^*, y-x\rangle)^2 + 2\Delta_S(x)\langle x^*, y-x\rangle = \Delta_S^2(y) - \Delta_S^2(x) \\ &\geq \|y-x\|^2 + 2\Delta_S(x)\Delta_S'(x; y-x) \geq \|y-x\|^2 + 2\Delta_S(x)\langle x^*, y-x\rangle, \end{aligned}$$

which is impossible.

(iv) Let $x^* \in \partial\Delta_S(x) \cap \mathbb{B}(0,1)$, $y^* \in \partial\Delta_S(y)$ with $\|x^*\| + \varepsilon \leq 1$ and $\Delta_S(x) + \Delta_S(y) \leq \eta\varepsilon$ for some $\varepsilon > 0$, $\eta > 0$. Suppose that $\|y-x\|^2 > \eta\langle y^* - x^*, y-x\rangle$, then

$$\Delta_S(y) - \Delta_S(x) \leq \langle y^*, y-x\rangle < \langle x^*, y-x\rangle + \frac{\|y-x\|^2}{\eta} \quad \text{and}$$

$$\begin{aligned} (\Delta_S(y) + \Delta_S(x))\left(\langle x^*, y-x\rangle + \frac{\|y-x\|^2}{\eta}\right) &> \Delta_S^2(y) - \Delta_S^2(x) \\ &\geq \|y-x\|^2 + 2\Delta_S(x)\langle x^*, y-x\rangle. \end{aligned}$$

Hence

$$(\Delta_S(y) - \Delta_S(x))\langle x^*, y-x\rangle + (\Delta_S(y) + \Delta_S(x))\frac{\|y-x\|^2}{\eta} > \|y-x\|^2$$

and consequently

$$\begin{aligned} \|x^*\|\,\|y-x\|^2 \geq (\Delta_S(y) - \Delta_S(x))\langle x^*, y-x\rangle &> \left(1 - \frac{\Delta_S(y) + \Delta_S(x)}{\eta}\right)\|y-x\|^2 \\ &\geq (1-\varepsilon)\|y-x\|^2, \end{aligned}$$

which is impossible. Thus (iv) holds true.

(v) Since Δ_S is a coercive convex function, so $\operatorname{argmin}\Delta_S \neq \emptyset$, see [1, Proposition 11.14] for details. Thus the set $\operatorname{argmin}\Delta_S$ contains at least one element. However it follows from (iii) that the set contains at most one element. □

4. Fréchet differentiability of the Fenchel conjugate of the antidistance function

Let us observe that if S is a singleton, say $S = \{s\}$, then for every $x^* \in \mathbb{B}[0,1]$ it holds $x^* \in \partial\Delta_S(s)$. Hence the operator $T : \mathbb{B}[0,1] \longrightarrow \mathbb{H}$ such that

$$\forall x^* \in \mathbb{B}(0,1), x^* \in \partial\Delta_S(\cdot)(T(x^*)) \tag{16}$$

is 0-Lipschitz, since $T(x^*) = s$ for all $x^* \in \mathbb{B}[0,1]$, where the notation

$$\partial\Delta_S(\cdot)(T(x^*)) := \partial\Delta_S(s) \text{ with } s = T(x^*),$$

is used. Thus the interesting case is whenever S is not a singleton. Throughout the rest of the presentation only this case is considered.

First of all let us recall that in [23, Propositon 6.3] it was shown that the operator

$$(\Delta_S)^{-1} : \mathbb{B}[0, 1-\varepsilon] \longrightarrow \mathbb{H}$$

is Lipschitz continuous for all $\varepsilon \in]0,1[$. Thus putting $T := (\Delta_S)^{-1}$ the existence of T, which is locally Lipschitz on $\mathbb{B}(0,1)$ and (16) holds true, is assured. Below another proof of this property is provided by means of the Perfect Square Criterion for Maximality.

Proposition 4.1. *Let $S \subset \mathbb{H}$ be a bounded subset. For each $x^* \in \mathbb{B}(0,1)$ there exists a unique $x \in \mathbb{H}$ such that $x^* \in \partial\Delta_S(x)$.*

Proof. Let us fix x^* such that $\|x^*\| < 1$ and for all $n \in \mathbb{N}$ put

$$M_n := \{(a, na^*) \in \mathbb{H} \times \mathbb{H} \mid a^* \in \partial\Delta_S(a)\}.$$

By (10) there are $w_n \in \mathbb{H}$, $w_n^* \in \mathbb{H}$ such that $(w_n, nw_n^*) \in M_n$ and

$$nx^* = w_n + nw_n^* \quad \text{for all } n \in \mathbb{N}. \tag{17}$$

Suppose that there is a subsequence $\{w_{n_k}\}_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \|w_{n_k}\| = \infty$. Since S is bounded, it follows from (15) that

$$1 = \lim_{k\to\infty} \frac{\|w_{n_k}\|^2 + \Delta_S(0)\|w_{n_k}\|}{\Delta_S(w_{n_k})\|w_{n_k}\|} \geq \lim_{k\to\infty} \frac{\langle w_{n_k}^*, w_{n_k}\rangle}{\|w_{n_k}\|} \geq \lim_{k\to\infty} \frac{\|w_{n_k}\|^2 - 2\Delta_S(0)\|w_{n_k}\|}{\Delta_S(w_{n_k})\|w_{n_k}\|} = 1.$$

Now by (17) we obtain

$$\begin{aligned} 0 &\leq \lim_{k\to\infty} \left(n_k\|x^*\|\|w_{n_k}\| - \|w_{n_k}\|^2 - n_k\langle w_{n_k}^*, w_{n_k}\rangle\right) \\ &= \lim_{k\to\infty} \|w_{n_k}\| \left(n_k\|x^*\| - \|w_{n_k}\| - n_k \frac{\langle w_{n_k}^*, w_{n_k}\rangle}{\|w_{n_k}\|}\right) = -\infty, \end{aligned}$$

a contradiction. Hence $\limsup_{n\to\infty} \|w_n\| < \infty$. Thus $\lim_{n\to\infty} \|w_n^* - x^*\| = 0$ by (17), which implies that any weak cluster point of the sequence $\{w_n\}_{n\in\mathbb{N}}$, let x be such a point, satisfies the following requirement

$$x^* \in \partial\Delta_S(x).$$

The uniqueness of the existence of such x follows from Corollary 3.2 (iii). □

A straightforward consequence of Proposition 4.1 is that there exists locally Lipschitz operator $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ such that

$$\forall x^* \in \mathbb{B}(0,1),\ x^* \in \partial\Delta_S(T(x^*)) \ \text{ and } \ T(\partial\Delta_S(T(x^*)) \cap \mathbb{B}(0,1)) = \{T(x^*)\}, \tag{18}$$

where the notations

$$\partial\Delta_S(\cdot)(T(x^*)) := \partial\Delta_S(x) \ \text{ with } \ x = T(x^*) \ \text{ and } \ \partial\Delta_S(T(x^*)) := \partial\Delta_S(\cdot)(T(x^*))$$

are used. Moreover the image of any weakly closed subset $C \subset \mathbb{B}[0, 1-\epsilon]$ under T is a bounded and closed subset of $\mathbb{B}[T(0), \frac{\Delta_S(T(0))}{2\varepsilon}]$, where $\varepsilon > 0$, which can be used to get additional information on points of the Fréchet differentiability of Δ_S.

Let us recall that by Proposition 3.1(iii) we have

$$\forall y \in \mathbb{H}, \partial\Delta_S(y) \subset \mathbb{S}[0,1] \Longrightarrow D_F\Delta_S(y) \text{ exists} . \tag{19}$$

Let us put

$$\mathbb{T} := \{\{t_i\}_{i\in\mathbb{N}} \mid \forall i \in \mathbb{N}, t_i > 0\} \tag{20}$$

and

$$\mathbb{D}(x, \{t_i\}_{i\in\mathbb{N}}) := \{h \in \mathbb{H} \mid \forall i \in \mathbb{N},\ \partial\Delta_S(x + t_i h) \subset \mathbb{S}[0,1]\} \tag{21}$$

for every $x \in \mathbb{H}$ and $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$.

Corollary 4.2. *The operator $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ which satisfies* (18) *has the following properties:*

(i) *it is locally Lipschitz on $\mathbb{B}(0,1)$;*

(ii) *for all $\varepsilon \in]0,1]$ and each weakly closed subset $C \subset \mathbb{B}[0,1-\epsilon]$, $C \neq \emptyset$ the image $T(C)$ is a bounded and closed subset such that*

$$T(C) \subset \begin{cases} \mathbb{B}[T(0), 2(1-\varepsilon)\Delta_S(T(0))], & \text{whenever } 2\varepsilon > 1; \\ \mathbb{B}[T(0), \frac{\Delta_S(T(0))}{2\varepsilon}], & \text{whenever } 2\varepsilon \leq 1; \end{cases}$$

(iii) *the set $\mathbb{D}(x, \{t_i\}_{i\in\mathbb{N}})$ is residual in $\mathbb{H}$ for each $x \in \mathbb{H}$ and $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$, that is, its complement is of first category, see* [16]. *Moreover, if a subset $\{t_1, t_2, \ldots\} \in \mathbb{T}$ is dense in $]0,\infty[$, then for all $h \in \mathbb{D}(x, \{t_i\}_{i\in\mathbb{N}})$ the subsets*

$$\{t \geq 0 \mid \partial\Delta_S(x + th) \subset \mathbb{S}[0,1]\}$$

are G_δ dense subsets of $]0,\infty[$.

Proof. (i) In order to show that T is locally Lipschitz on $\mathbb{B}(0,1)$ it is enough to show that the set $T(\mathbb{B}[0,1-\varepsilon])$ is bounded for each $\varepsilon \in]0,1[$, see Corollary 3.2(iv). Let us fix $\varepsilon \in]0,1[$ and notice that $\{T(0)\} = \operatorname{argmin} \Delta_S$ by Corollary 3.2(v). For every $x^* \in \mathbb{B}[0,1-\varepsilon]$ we have $\Delta_S(T(x^*)) \geq \Delta_S(T(0))$ and

$$\begin{aligned} \Delta_S^2(T(0)) &\geq \Delta_S^2(T(x^*)) + 2\Delta_S(T(x^*))\langle x^*, T(0) - T(x^*)\rangle + \|T(0) - T(x^*))\|^2 \\ &\geq \Delta_S^2(T(x^*)) - 2\Delta_S(T(x^*))(1-\varepsilon)\|T(0) - T(x^*)\| + \|T(0) - T(x^*))\|^2 \\ &= (\Delta_S(T(x^*)) - \|T(0) - T(x^*)\|)^2 + 2\varepsilon\Delta_S(T(x^*))\|T(x^*) - T(0)\| \\ &\geq 2\varepsilon\Delta_S(T(0))\|T(x^*) - T(0)\|. \end{aligned} \tag{22}$$

Thus $T(\mathbb{B}[0,1-\varepsilon]) \subset \mathbb{B}[T(0), \frac{\Delta_S(T(0))}{2\varepsilon}]$, so it is a bounded subset and hence T is locally Lipschitz.

(ii) Assume that $\varepsilon \in]0,1]$ and $C \subset \mathbb{B}[0,1-\epsilon]$, $C \neq \emptyset$ is a weakly closed subset. Let $\{c_i\}_{i\in\mathbb{N}}$ be a sequence such that $c_i \in C$ for all $i \in \mathbb{N}$ and $x = \lim_{i\to\infty} T(c_i)$. Let $c \in C$ be a weak sequential cluster point of the sequence $\{c_i\}_{i\in\mathbb{N}}$. Since

$$\forall i \in \mathbb{N},\ c_i \in \partial\Delta_S(\cdot)(T(c_i)) \ \text{ and } \ x = \lim_{i\to\infty} T(c_i),$$

so $c \in \partial\Delta_S(x) \cap C$, thus $x = T(c) \in T(C)$.

In order to terminate the proof of (ii) let us fix $\varepsilon \in]\frac{1}{2}, 1]$. Observe that

$$\Delta_S^2(T(x^*)) - \Delta_S^2(T(0)) \geq \|T(x^*) - T(0)\|^2$$

by Corollary 3.2(iv). Thus

$$\Delta_S(T(x^*)) + \Delta_S(T(0)) \geq 2\|T(x^*) - T(0)\|,$$

so $$(\Delta_S(T(x^*)) - \|T(0) - T(x^*)\|)^2 > (\Delta_S(T(0)) - \|T(0) - T(x^*)\|)^2$$

in this case. It follows from (22) that

$$\begin{aligned}\Delta_S^2(T(0)) &\geq (\Delta_S(T(x^*)) - \|T(0) - T(x^*)\|)^2 + 2\varepsilon\Delta_S(T(x^*))\|T(x^*) - T(0)\| \\ &\geq (\Delta_S(T(0)) - \|T(0) - T(x^*)\|)^2 + 2\varepsilon\Delta_S(T(0))\|T(x^*) - T(0)\|,\end{aligned}$$

which justifies the inclusion $T(C) \subset \mathbb{B}[T(0), 2(1-\varepsilon)\Delta_S(T(0))]$.

(iii) Let us fix $x \in \mathbb{H}$ and $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$. The set

$$\left(\bigcup_{k\in\{1,\ldots,i\}} t_k^{-1}(T(B[0, 1 - i^{-1}]) - x)\right)$$

is closed for each $i \in \mathbb{N} \setminus \{1\}$, see part (ii) of the proof. Since Δ_S is Fréchet differentiable on a dense subset of $\mathbb{H}$, see for example [17, Theorem 2.34], we conclude that $\operatorname{int}(T(B[0, 1 - i^{-1}]) = \emptyset$ for all $i \in \mathbb{N} \setminus \{1\}$. Thus the set

$$\bigcup_{i\in\mathbb{N}\setminus\{1\}} \left(\bigcup_{k\in\{1,\ldots,i\}} t_k^{-1}(T(B[0, 1 - i^{-1}]) - x)\right)$$

is of first category, see [16, p. 40]. It is easy to notice that

$$\mathbb{H} \setminus \mathbb{D}(x, \{t_i\}_{i\in\mathbb{N}}) \subset \bigcup_{i\in\mathbb{N}\setminus\{1\}} \left(\bigcup_{k\in\{1,\ldots,i\}} t_k^{-1}(T(\mathbb{B}[0, 1 - i^{-1}]) - x)\right),$$

which justifies the first part of the statement in (iii), see for example [21, Baire Category, p. 43(a)]. In order to terminate the proof let us fix a dense subset $\{t_1, t_2, \ldots\}$ of $]0, \infty[$, then for all $h \in \mathbb{D}(x, \{t_i\}_{i\in\mathbb{N}})$ and $n \in \mathbb{N}$ we have the following inclusion

$$\{t_1, t_2, \ldots\} \subset U_n := \{t > 0 \mid x + th \notin T(B[0, 1 - n^{-1}]\},$$

hence $\bigcap_{n\in N} U_n$ is a G_δ dense subset of $]0, \infty[$. Since the equality

$$\{t \geq 0 \mid \partial\Delta_S(x + th) \subset \mathbb{S}[0, 1]\} = \bigcap_{n\in N} U_n$$

is easy to observe, so the second part of the statement in (iii) is justified. □

Fact 4.3. *Let $T : \mathbb{B}(0, 1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18). *Then for every $x^* \in \mathbb{B}(0, 1)$ and $h^* \in \mathbb{H}$ such that $\|x^* + h^*\| < 1$ the following estimation*

$$0 \leq \Delta_S(T(x^*+h^*)) - \Delta_S(T(x^*)) - \langle x^*, T(x^*+h^*) - T(x^*)\rangle \leq \langle h^*, T(x^*+h^*) - T(x^*)\rangle$$

holds true.

Proof. Let us fix $x^* \in \mathbb{H}$ and $h^* \in \mathbb{H}$ such that $\|x^*\| < 1$ and $\|x^* + h^*\| < 1$. Since $x^* \in \partial\Delta_S(T(x^*))$ and $x^* + h^* \in \partial\Delta_S(T(x^* + h^*))$, so

$$\begin{aligned} 0 &\leq \Delta_S(T(x^* + h^*)) - \Delta_S(T(x^*)) - \langle x^*, T(x^* + h^*) - T(x^*)\rangle \\ &\leq \langle x^* + h^*, T(x^* + h^*) - T(x^*)\rangle - \langle x^*, T(x^* + h^*) - T(x^*)\rangle \end{aligned}$$

which justifies the inequality. □

For fixed $x \in \mathbb{H}$ and $\varepsilon \in]0, 1[$ a condition allowing to identify the direction $v \in \mathbb{H}$ such that $\Delta'_S(x; v) = \langle y^*, v\rangle$ for some $y^* \in \partial\Delta_S(x) \cap \mathbb{B}[0, 1-\varepsilon]$ is given in the next Corollary.

Corollary 4.4. *Let $T : \mathbb{B}(0, 1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18) *and $\varepsilon \in]0, 1[$ be given. Then*

$$\exists y^* \in \partial\Delta_S(x) \cap \mathbb{B}[0, 1-\varepsilon], \quad \Delta'_S(x; v) = \langle y^*, v\rangle \quad \textit{and} \quad x = T(y^*)$$

for all $x \in \mathbb{H}$ and $v \in \mathbb{H}$ for which there is a sequence $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$, see (20), *converging to zero such that*

$$\lim_{i\to\infty} \frac{d_{T(\mathbb{B}[0,1-\varepsilon])}(x + t_i v)}{t_i} = 0.$$

Moreover, if $x \in \mathbb{H}$ and $v \in \mathbb{H}$ are such that

$$\lim_{i\to\infty} \frac{d_{T(\mathbb{B}[0,1-\varepsilon])}(x + t_i v)}{t_i} > 0$$

for some $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$ converging to zero, then

$$\forall y^* \in \partial\Delta_S(x), \Delta'_S(x; v) = \langle y^*, v\rangle \Longrightarrow \|y^*\| \geq 1 - \varepsilon.$$

Proof. Let $x \in \mathbb{H}$ and $v \in \mathbb{H}$ be such that there is a sequence $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$ converging to zero for which

$$\lim_{i\to\infty} \frac{d_{T(\mathbb{B}[0,1-\varepsilon])}(x + t_i v)}{t_i} = 0.$$

For each $i \in \mathbb{N}$ there is $y_i^* \in \mathbb{B}[0, 1-\varepsilon]$ such that

$$\lim_{i\to\infty} \frac{\|x + t_i v - T(y_i^*)\|}{t_i} = 0.$$

Without loss of generality we may assume that $y_i^* \overset{weak}{\longrightarrow} y^* \in \mathbb{B}[0, 1-\varepsilon]$, therefore $y^* \in \partial\Delta_S(x) \cap \mathbb{B}[0, 1-\varepsilon]$ and

$$\begin{aligned} \Delta'_S(x; v) &= \lim_{i\to\infty} \frac{\Delta_S(x + t_i v) - \Delta_S(x)}{t_i} = \lim_{i\to\infty} \frac{\Delta_S(T(y_i^*)) - \Delta_S(T(y^*))}{t_i} \\ &= \lim_{i\to\infty} \frac{\langle y_i^*, T(y_i^*) - T(y^*)\rangle}{t_i} = \langle y^*, v\rangle, \end{aligned}$$

and the first part of the statement is justified.

In order to terminate the proof let us consider $x \in \mathbb{H}$ and $v \in \mathbb{H}$ are such that

$$\lim_{i\to\infty} \frac{d_{T(\mathbb{B}[0,1-\varepsilon])}(x+t_i v)}{t_i} > 0 \tag{23}$$

for some sequence $\{t_i\}_{i\in\mathbb{N}} \in \mathbb{T}$ converging to zero.

Suppose that $\exists\, y^* \in \partial\Delta_S(x),\ \ \Delta_S'(x;v) = \langle y^*, v\rangle$ and $\|y^*\| < 1-\varepsilon$.

Put
$$\varepsilon(t) := \sqrt{\frac{\Delta_S(x+tv) - \Delta_S(x) - \langle y^*, tv\rangle}{t}},$$

for all $t \in]0,\infty[$. If $\varepsilon(\bar{t}) = 0$ for some $\bar{t} > 0$, then $y^* \in \partial\Delta_S(x+tv)$ for all $t \in]0,\bar{t}]$ and consequently

$$\lim_{i\to\infty} \frac{d_{T(\mathbb{B}[0,1-\varepsilon])}(x+t_i v)}{t_i} = 0,$$

which contradicts (23). We have

$$\forall t \in]0,\infty[,\quad 0 \le \Delta_S(x+tv) - \Delta_S(x) - \langle y^*, tv\rangle \le t(\varepsilon(t))^2$$

and
$$\forall a \in \mathbb{H},\quad 0 \le \Delta_S(x+a) - \Delta_S(x) - \langle y^*, a\rangle.$$

By Theorem 2.1 there are $a(t) \in \mathbb{B}[x+tv, t\varepsilon(t)]$ and $a^*(t) \in \partial\Delta_S(a(t)) \cap \mathbb{B}[y^*, \varepsilon(t)]$, so

$$T(y^* + (a^*(t) - y^*)) - (x+tv) = a(t) - (x+tv) \in \mathbb{B}[0, t\varepsilon(t)]$$

for all $t \in]0,\infty[$ small enough. Thus, $a(t) \in T(\mathbb{B}[0,1-\varepsilon])$ and

$$d_{T(\mathbb{B}[0,1-\varepsilon])}(x+tv) \le \|a(t) - (x+tv)\| \le t\varepsilon(t),$$

for all $t > 0$ small enough, which contradicts (23). Thus the second part of the statement holds true. □

It is known that for every $x \in \mathbb{H}$ and $v \in \mathbb{H}$ there is $v^* \in \partial\Delta_S(x)$ such that

$$\Delta_S'(x;v) = \langle v^*, v\rangle, \tag{24}$$

see [1, Theorem 17.19]. Thus, the following problem arises: if $x = T(x^*)$, then find solutions of the following equations

$$\Delta_S'(T(x^*);v) = \langle x^*, v\rangle. \tag{25}$$

In order to characterize solutions of the equation in (25) for $T : B(0,1) \longrightarrow \mathbb{H}$ being an operator satisfying (18) and $x^* \in \mathbb{B}(0,1)$, $f \in \mathscr{F}$, $g \in \mathscr{G}$, see (1) and (2), and $t \ge 0$, let us put

$$D(T,f,g,x^*)(t) := \left\{ h^* \in \mathbb{H} \;\middle|\; \begin{array}{l} \|x^* + h^*\| < 1,\ h^* \in \mathbb{B}[0, f(t)] \\ \text{and } T(x^*+h^*) \in \mathbb{B}[T(x^*), tg(t)] \end{array} \right\}, \tag{26}$$

where $\mathbb{B}[c,\infty] := \mathbb{H}$ for all $c \in \mathbb{H}$.

Below, see Proposition 4.5, it is shown that if for some $f \in \mathscr{F}$, $g \in \mathscr{G}$

$$v \in \bigcap_{\delta>0} \operatorname{cl}\operatorname{conv}\Big(\bigcup_{\substack{t\in]0,\delta],\\ h^*\in D(T,f,g,x^*)(t)}} \{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\}\Big), \tag{27}$$

then (25) holds true. On the other hand if $\|x^*\|<1$ and (25) is valid, then (27) is satisfied for some $f \in \mathscr{F}$, $g \in \mathscr{G}$, see Proposition 4.6.

Proposition 4.5. *Let* $T:\mathbb{B}(0,1)\longrightarrow\mathbb{H}$ *be an operator which satisfies* (18). *Then for every* $x^*\in\mathbb{B}(0,1)$ *and* $f\in\mathscr{F}$, $g\in\mathscr{G}$ *the following property holds true*

$$\forall v\in \bigcap_{\delta>0} \operatorname{cl}\operatorname{conv}\Big(\bigcup_{\substack{t\in]0,\delta],\\ h^*\in D(T,f,g,x^*)(t)}} \{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\}\Big), \quad \Delta_S'(T(x^*);v)=\langle x^*,v\rangle.$$

Proof. Fix some $x^*\in\mathbb{B}(0,1)$, $f\in\mathscr{F}$, $g\in\mathscr{G}$ and

$$v \in \bigcap_{\delta>0} \operatorname{cl}\operatorname{conv}\Big(\bigcup_{t\in]0,\delta], h^*\in D(T,f,g,x^*)(t)} \{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\}\Big).$$

There are sequences $\{\delta_i\}_{i\in\mathbb{N}}$, $\{k_i\}_{i\in\mathbb{N}}$ such that $\delta_i\searrow 0$, $k_i\in\mathbb{N}$, $\lambda_i^j>0$, $t_i^j\in]0,\delta_i]$, $h_i^j\in D(T,f,g,x^*)(t_i^j)$ for all $i\in\mathbb{N}$ and $j\in\{1,\dots,k_i\}$ with $\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j=1$ for all $i\in\mathbb{N}$, and

$$v=\lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\left(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\right).$$

It follows from the convexity of $\Delta_S(\cdot)$ and Fact 4.3 that

$$\begin{aligned}
\langle x^*,v\rangle &\le \Delta_S'(T(x^*);v)=\Delta_S'\Big(T(x^*);\lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\Big(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\Big)\Big)\\
&\le \lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\Delta_S'\Big(T(x^*);\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\Big)\\
&= \lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\lim_{s\searrow 0}\frac{\Delta_S\Big(T(x^*)+s\left(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\right)\Big)-\Delta_S(T(x^*))}{s}\\
&\le \lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\frac{\Delta_S(T(x^*+h_i^j))-\Delta_S(T(x^*))}{t_i^j}\\
&\le \lim_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\langle x^*+h_i^j,\left(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\right)\rangle\\
&\le \lim_{i\to\infty}\langle x^*,\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\left(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\right)\rangle+\limsup_{i\to\infty}\sum_{j\in\{1,\dots,k_i\}}\lambda_i^jf(t_i^j)g(t_i^j)\\
&\le \lim_{i\to\infty}\langle x^*,\sum_{j\in\{1,\dots,k_i\}}\lambda_i^j\left(\frac{1}{t_i^j}(T(x^*+h_i^j)-T(x^*))\right)\rangle=\langle x^*,v\rangle,
\end{aligned}$$

which justifies the statement. □

Proposition 4.6. *Let $T\colon \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18). *Then for every $x^* \in \mathbb{B}(0,1)$ and $v \in \mathbb{H}$ such that* (25) *is fulfilled there are $f \in \mathscr{F}$, $g \in \mathscr{G}$ such that*

$$v \in \bigcap_{\delta>0}\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{\substack{t\in]0,\delta],\\ h^*\in D(T,f,g,x^*)(t)}} \left\{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\right\}\right)\right).$$

Proof. Let $x^* \in \mathbb{B}(0,1)$, $x = T(x^*)$ and $v \in \mathbb{H}$ be such that (25) is fulfilled. Put $f(0) := 0$ and for all $t \in]0,\infty[$ define

$$\varepsilon(t) := \sqrt{\frac{\Delta_S(x+tv)-\Delta_S(x)-\langle x^*, tv\rangle}{t}},$$

$$g(t) := \|v\| + 1 + \varepsilon(t) \quad \text{and} \quad f(t) := \begin{cases} t+\varepsilon(t), & \text{if } t>0;\\ 0, & \text{if } t=0.\end{cases}$$

Assume that $\varepsilon(\bar{t}) = 0$ for some $\bar{t} > 0$. Then $x^* \in \partial\Delta_S(x+tv)$ for all $t \in]0,\bar{t}]$. Thus it follows from Corollary 3.2 (iii) that $v = 0$, so

$$v \in \bigcap_{\delta>0}\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{t\in]0,\delta],h^*\in D(T,f,g,x^*)(t)} \left\{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\right\}\right)\right).$$

Let us consider the case $v \neq 0$. By (25) we get

$$\forall t \in]0,\infty[, \quad 0 \le \Delta_S(x+tv)-\Delta_S(x)-\langle x^*, tv\rangle \le t(\varepsilon(t))^2$$

and
$$\forall a \in \mathbb{H}, \quad 0 \le \Delta_S(x+a)-\Delta_S(x)-\langle x^*, a\rangle.$$

By Theorem 2.1 there are $a(t)\in\mathbb{B}[x+tv, t\varepsilon(t)]$ and $a^*(t)\in\partial\Delta_S(a(t))\cap\mathbb{B}[x^*,\varepsilon(t)]$, so

$$T(x^*+(a^*(t)-x^*))-T(x^*) = tv + (a(t)-(x+tv)) \in \mathbb{B}[0, tg(t)],$$

for all $t \in]0,\infty[$. Thus, since $f \in \mathscr{F}$, $g \in \mathscr{G}$, we get

$$v \in \bigcap_{\delta>0}\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{\substack{t\in]0,\delta],\\ h^*\in D(T,f,g,x^*)(t)}} \left\{\tfrac{1}{t}(T(x^*+h^*)-T(x^*))\right\}\right)\right)$$

and the statement holds true. $\square$

Corollary 4.7. *Let $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18). *Then for every $x^* \in \mathbb{B}(0,1)$, $f \in \mathscr{F}$, $g \in \mathscr{G}$ and $y^* \in \partial\Delta_S(T(x^*))$ the following property*

$$\forall v \in \mathbb{H}, \quad \{-v, v\} \subset \bigcap_{\delta>0}\left(\operatorname{cl}\operatorname{conv}\left(\bigcup_{\substack{t\in]0,\delta],\\ h^*\in D(T,f,g,x^*)(t)}} \left\{\frac{1}{t}(T(x^*+h^*)-T(x^*))\right\}\right)\right)$$
$$\Longrightarrow \langle y^*, v\rangle = \langle x^*, v\rangle = \Delta'_S(T(x^*); v)$$

holds true.

Proof. Let us fix $x^* \in \mathbb{B}(0,1)$, $f \in \mathscr{F}$, $g \in \mathscr{G}$, $y^* \in \Delta_S(T(x^*))$ and $v \in \mathbb{H}$ such that we have

$$\{-v, v\} \subset \bigcap_{\delta>0} \left(\operatorname{cl}\operatorname{conv} \left(\bigcup_{\substack{t \in]0,\delta], \\ h^* \in D(T,f,g,x^*)(t)}} \left\{ \frac{1}{t}(T(x^*+h^*) - T(x^*)) \right\} \right) \right).$$

By Proposition 4.5 we obtain

$$\langle y^*, v \rangle \le \Delta'_S(T(x^*); v) = \langle x^*, v \rangle \text{ and } \langle y^*, -v \rangle \le \Delta'_S(T(x^*); -v) = \langle x^*, -v \rangle,$$

which implies the desired equalities. □

A simple consequence of Proposition 4.6 is the following

Fact 4.8. *Let $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18)*. Then for every $x^* \in \mathbb{B}(0,1)$ and $v \in \mathbb{H}$ if*

$$\forall y^* \in \partial \Delta_S(T(x^*)), \quad \Delta'_S(T(x^*); v) = \langle y^*, v \rangle,$$

then for some $f \in \mathscr{F}$ and $g \in \mathscr{G}$ the following property

$$\{-v, v\} \subset \bigcap_{\delta>0} \left(\operatorname{cl}\operatorname{conv} \left(\bigcup_{\substack{t \in]0,\delta], \\ h^* \in D(T,f,g,x^*)(t)}} \left\{ \frac{1}{t}(T(x^*+h^*) - T(x^*)) \right\} \right) \right)$$

holds true.

Proof. Let us fix $x^* \in \mathbb{B}(0,1)$ and $v \in \mathbb{H}$ such that

$$\forall y^* \in \partial \Delta_S(T(x^*)), \quad \Delta'_S(T(x^*); v) = \langle y^*, v \rangle.$$

The convexity of $\Delta'_S(T(x^*); \cdot)$ yields

$$\forall y^* \in \partial \Delta_S(T(x^*)), \quad \Delta'_S(T(x^*); -v) = \langle y^*, -v \rangle.$$

It follows from Proposition 4.6 there are $f_i \in \mathscr{F}$, $g_i \in \mathscr{G}$ for $i \in \{1, 2\}$ such that

$$v \in \bigcap_{\delta>0} \left(\operatorname{cl}\operatorname{conv} \left(\bigcup_{\substack{t \in]0,\delta], \\ h^* \in D(T,f_1,g_1,x^*)(t)}} \left\{ \frac{1}{t}(T(x^*+h^*) - T(x^*)) \right\} \right) \right).$$

and

$$-v \in \bigcap_{\delta>0} \left(\operatorname{cl}\operatorname{conv} \left(\bigcup_{\substack{t \in]0,\delta], \\ h^* \in D(T,f_2,g_2,x^*)(t)}} \left\{ \frac{1}{t}(T(x^*+h^*) - T(x^*)) \right\} \right) \right).$$

Hence for $f := \max\{f_1, f_2\}$ and $g := \max\{g_1, g_2\}$ we get

$$\{-v, v\} \subset \bigcap_{\delta>0} \left(\operatorname{cl}\operatorname{conv} \left(\bigcup_{\substack{t \in]0,\delta], \\ h^* \in D(T,f,g,x^*)(t)}} \left\{ \frac{1}{t}(T(x^*+h^*) - T(x^*)) \right\} \right) \right). \qquad \square$$

We recall that the Fenchel conjugate of Δ_S is denoted by Δ_S^*. In order to prove the Fréchet differeniability of Δ_S^* we use the following property of subdifferential mapping, see [10, Corollary 4.2]:

> *if the subdifferential mapping ∂f of a convex function f has a continuous selection on an open convex subset $U \subset \mathbb{H}$, that is $\sigma(u) \in \partial f(u)$ for all $u \in U$ and σ is a continuous operator, then the Fréchet derivative $D_F f(u)$ exists for all $u \in U$ and $\sigma(u) = D_F f(u)$ for all $u \in U$.*

Theorem 4.9. *Let $S \subset \mathbb{H}$ be a bounded subset. Then the Fréchet derivative of Δ_S^* exists at each $x^* \in \mathbb{B}(0,1)$ and $D_F \Delta_S^*(x^*) = T(x^*)$, where $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ is such that* (18) *holds true. Hence the Fréchet derivative of Δ_S^* is a locally Lipschitz operator on* $\mathbb{B}(0,1)$.

Proof. For all $x^* \in \mathbb{B}(0,1)$ by (18) and (9) we have

$$\Delta_S(T(x^*)) + \Delta_S^*(x^*) = \langle x^*, T(x^*)\rangle.$$

It follows from Corollary 4.2 that T is locally Lipschitz on $\mathbb{B}(0,1)$. In order to terminate the proof it is enough to observe that it follows from (9) and (18), see also [1, Theorem 16.23](iv), that

$$\forall x^* \in \mathbb{B}(0,1), T(x^*) \in \partial\Delta_S^*(x^*),$$

so T is a continuous selection of $\partial\Delta_S^*$ on $\mathbb{B}(0,1)$ and the property mentioned just before this Theorem justifies the statement. □

5. Some properties of derivatives of T

Second order differentiability is a useful tool in investigating optimality problems. However being beyond the classical framework we encounter formidable challenges in exploring second order properties of functions under investigation. A lot of information on the second order theory of semidifferentiability in the finite dimensional setting can be found in [18, Chapter 13, Second-Order Theory]; a study of the second order behavior of convex functions over a reflexive Banach space can be found in [15] for example. However we know little whenever the infinite dimensional setting is under consideration. Most questions concerning the second order theory of differentiability are open in this case, see for example [26]. Perhaps, it is a reason why some old questions are still open, see [9]. The progress in the developing this theory is modest, compare for example [2] and [5], see also [26].

In this section some properties of a derivative of the operator T, which existence is preserved in Proposition 4.1, see (18), are explored. It is known that the second order differeniability of Δ_S^* is preserved at points where differentiability of T is guaranteed, see [3, Theorem 3.1] for example. The existence of such points can be inferred from the classic result concerning the Gâteaux differentiability of a Lipschitz mapping on a separable Hilbert space, see [14], see also [13] for some recent information on differentiability of Lipschitz operators.

We recall that the notations

$$\partial\Delta_S(\cdot)(y) := \partial\Delta_S(y), \quad \partial\Delta_S(T(x^*)) := \partial\Delta_S(\cdot)(T(x^*)),$$

$$\ker A := \{w^* \in \mathbb{H} \mid Aw^* = 0\} \ \text{ and } \ A(\mathbb{H}) := \{y \in \mathbb{H} \mid \exists x \in \mathbb{H} : y = Ax\}$$

are used, whenever $A : \mathbb{H} \longrightarrow \mathbb{H}$ is a linear operator.

A first observation concerns additivity of the directional derivative of Δ_S on some elements.

Corollary 5.1. *Let* $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ *be an operator satisfying* (18), $x^* \in B(0,1)$ *and let* $A\colon \mathbb{H} \longrightarrow \mathbb{H}$ *be a bounded linear operator such that*

$$\forall h^* \in \mathbb{H}, \quad \lim_{t\to 0} \frac{\|T(x^*+th^*) - T(x^*) - tAh^*\|}{t} = 0,$$

then $\forall h \in \mathbb{H}, \ \forall v^* \in \mathbb{H}:$

$$\Delta'_S(T(x^*); h+Av^*) = \Delta'_S(T(x^*); h) + \Delta'_S(T(x^*); Av^*) = \Delta'_S(T(x^*); h) + \langle x^*, Av^*\rangle.$$

Moreover, if $\forall h \in \ker A, \ \Delta'_S(T(x^*); h) = \langle x^*, h\rangle$, *then* $Q_S(T(x^*)) = \emptyset$, *see* (12).

Proof. It follows from Corollary 4.7 that for all $y^* \in \partial\Delta_S(T(x^*))$ and $v^* \in \mathbb{H}$ the following property holds true

$$\langle y^*, Av^*\rangle = \Delta'_S(T(x^*); Av^*).$$

In fact, let us fix $v^* \in \mathbb{H}\setminus\{0\}$. For $t \geq 0$ put $f(t) := \|v^*\|t$ and

$$g(t) := \begin{cases} 1 + \frac{\|T(x^*+tv^*)-T(x^*)\|+\|T(x^*-tv^*)-T(x^*)\|}{t}, & \text{if } \max\{\|x^*+tv^*\|, \|x^*-tv^*\|\} < 1, \\ \infty, & \text{if } 1 \leq \max\{\|x^*+tv^*\|, \|x^*-tv^*\|\}. \end{cases}$$

It is easy to observe, see (26), that we have

$$\begin{aligned} Av^* &\in \bigcap_{\delta>0} \mathrm{cl}\left(\bigcup_{t\in]0,\delta]} \left\{\tfrac{1}{t}(T(x^*+tv^*) - T(x^*))\right\}\right) \\ &\subset \bigcap_{\delta>0} \mathrm{cl}\,\mathrm{conv}\left(\bigcup_{\substack{t\in]0,\delta],\\ w^*\in D(T,f,g,x^*)(t)}} \left\{\tfrac{1}{t}(T(x^*+w^*) - T(x^*))\right\}\right), \end{aligned}$$

which by Proposition 4.5 establishes the equality $\Delta'_S(T(x^*); Av^*) = \langle x^*, Av^*\rangle$. Since $\langle x^*, A(-v^*)\rangle = \langle x^*, -Av^*\rangle$, so repeating the reasoning above, we have

$$\begin{aligned} -Av^* &\in \bigcap_{\delta>0} \mathrm{cl}\left(\bigcup_{t\in]0,\delta]} \left\{\tfrac{1}{t}(T(x^*-tv^*) - T(x^*))\right\}\right) \\ &\subset \bigcap_{\delta>0} \mathrm{cl}\,\mathrm{conv}\left(\bigcup_{\substack{t\in]0,\delta],\\ w^*\in D(T,f,g,x^*)(t)}} \left\{\tfrac{1}{t}(T(x^*+w^*) - T(x^*))\right\}\right), \end{aligned}$$

and $\langle x^*, A(-v^*)\rangle = \Delta'_S(T(x^*); -Av^*)$. Hence it follows from Corollary 4.7 that $\langle y^*, Av^*\rangle = \langle x^*, Av^*\rangle = \Delta'_S(T(x^*); Av^*)$ for all $y^* \in \partial\Delta_S(T(x^*))$. Thus

$$\begin{aligned} \Delta'_S(T(x^*); h+Av^*) &= \sup_{y^*\in\partial\Delta_S(T(x^*))} \langle y^*, h+Av^*\rangle = \sup_{y^*\in\partial\Delta_S(T(x^*))} \langle y^*, h\rangle + \langle y^*, Av^*\rangle \\ &= \sup_{y^*\in\partial\Delta_S(T(x^*))} \langle y^*, h\rangle + \Delta'_S(T(x^*); Av^*) = \Delta'_S(T(x^*); h) + \Delta'_S(T(x^*); Av^*), \end{aligned}$$

which justifies the first statement.

Assume additionally that $\forall h \in \ker A,\ \Delta'_S(T(x^*); h) = \langle x^*, h\rangle$.

It follows from Fact 2.3 that $\mathbb{H} = \ker A + \operatorname{cl} A(\mathbb{H})$, so using the additional assumption and the first part of the proof we obtain

$$\forall h \in \mathbb{H}, \quad \Delta'_S(T(x^*); h) = \langle x^*, h\rangle.$$

Hence, if $Q_S(T(x^*)) \neq \emptyset$, then $\partial\Delta_S(T(x^*)) \subset \mathbb{S}[0,1]$, so by Proposition 3.1 we obtain the following contradiction $D_F\Delta_S(T(x^*))$ exists and $x^* \in \partial\Delta_S(T(x^*))$. □

The symmetry of the second order derivative is expected property. Such a result can be obtained even for nonsmooth functions, see for example [19]. Below it is shown that Gâteaux derivative of the Fréchet derivative of the Fenchel conjugate of antidistance function is a symmetric operator.

Fact 5.2. *Let $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18)*, $x^* \in B(0,1)$ and let $A : \mathbb{H} \longrightarrow \mathbb{H}$ be a bounded linear operator such that*

$$\forall h^* \in \mathbb{H}, \quad \lim_{t\to 0} \frac{\|T(x^* + th^*) - T(x^*) - tAh^*\|}{t} = 0,$$

then $\forall h^* \in \mathbb{H}$, $v^* \in \mathbb{H}$ we have $\langle v^*, Ah^*\rangle = \langle Av^*, h^*\rangle$.

Proof. If $h^* = 0$ or $v^* = 0$ then $0 = \langle v^*, Ah^*\rangle = \langle Av^*, h^*\rangle$.

Thus the statement is justified in this case. Let us notice that in order to get the equality in general case it is enough to establish it for any $h^* \in \mathbb{H}$ and $v^* \in \mathbb{H}$ such that $1 = \|h^*\| = \|v^*\|$. For every $t \in [0, \frac{1-\|x^*\|}{\|h^*\|+\|v^*\|}[$ let us define the following function $f_t : [0,t] \longrightarrow \mathbb{R}$ as follows

$$\begin{aligned}\forall s \in [0,t], f_t(s) := {} & \Delta^*_S(x^* + th^* + sv^*) - \Delta^*_S(x^* + th^*) - \Delta^*_S(x^* + sv^*) \\ & - \Delta^*_S(x^* + sh^* + tv^*) + \Delta^*_S(x^* + sh^*) + \Delta^*_S(x^* + tv^*).\end{aligned}$$

Let us notice that $f_t(0) = f_t(t) = 0$ for all $t \in [0, \frac{1-\|x^*\|}{\|h^*|+\|v^*\|}[$, thus by the classical Mean Value Theorem there are $\theta(t) \in]0,t[$ such that

$$\begin{aligned} 0 &= \langle T(x^* + th^* + \theta(t)v^*), v^*\rangle - \langle T(x^* + \theta(t)v^*), v^*\rangle \\ &\quad - \langle T(x^* + \theta(t)h^* + tv^*), h^*\rangle + \langle T(x^* + \theta(t)h^*), h^*\rangle \end{aligned}$$

for all $t \in [0, \frac{1-\|x^*\|}{\|h^*|+\|v^*\|}[$, see for example [20, Chapter 5, Mean Value Theorems]. Thus for all $t \in [0, \frac{1-\|x^*\|}{\|h^*\|+\|v^*\|}[$, using the assumption $1 = \|h^*\| = \|v^*\|$, it follows that

$$\begin{aligned} & \frac{\|T(x^* + th^* + \theta(t)v^*) - T(x^* + \theta(t)v^*) - tAh^*\|}{t} \\ & \quad + \frac{\|T(x^* + \theta(t)h^* + tv^*) - T(x^* + \theta(t)h^*) - tAv^*\|}{t} \\ & \geq \frac{\langle T(x^* + th^* + \theta(t)v^*) - T(x^* + \theta(t)v^*) - tAh^*, -v^*\rangle}{t} \\ & \quad + \frac{\langle T(x^* + \theta(t)h^* + tv^*) - T(x^* + \theta(t)h^*) - tAv^*, h^*\rangle}{t} = \langle Ah^*, v^*\rangle - \langle Av^*, h^*\rangle \end{aligned}$$

and

$$\frac{\|T(x^*+th^*+\theta(t)v^*)-T(x^*+\theta(t)v^*)-tAh^*\|}{t}$$
$$+\frac{\|T(x^*+\theta(t)h^*+tv^*)-T(x^*+\theta(t)h^*)-tAv^*\|}{t}$$
$$\geq \frac{\langle T(x^*+th^*+\theta(t)v^*)-T(x^*+\theta(t)v^*)-tAh^*,v^*\rangle}{t}$$
$$+\frac{\langle T(x^*+\theta(t)h^*+tv^*)-T(x^*+\theta(t)h^*)-tAv^*,-h^*\rangle}{t}=-\langle Ah^*,v^*\rangle+\langle Av^*,h^*\rangle,$$

hence

$$\langle Ah^*,v^*\rangle-\langle Av^*,h^*\rangle-\frac{\|T(x^*+th^*+\theta(t)v^*)-T(x^*+\theta(t)v^*)-tAh^*\|}{t}$$
$$-\frac{\|T(x^*+\theta(t)h^*+tv^*)-T(x^*+\theta(t)h^*)-tAv^*\|}{t}\leq 0$$
$$\leq \langle Ah^*,v^*\rangle-\langle Av^*,h^*\rangle+\frac{\|T(x^*+th^*+\theta(t)v^*)-T(x^*+\theta(t)v^*)-tAh^*\|}{t}$$
$$+\frac{\|T(x^*+\theta(t)h^*+tv^*)-T(x^*+\theta(t)h^*)-tAv^*\|}{t}.$$

Now the equality $0=\langle Ah^*,v^*\rangle-\langle Av^*,h^*\rangle$ follows after passage to the limit as $t\downarrow 0$. $\square$

A straightforward consequence of Proposition 4.5 is the following property of the Gâteaux derivative of the Fréchet derivative of the Fenchel conjugate of antidistance function.

Fact 5.3. *Let $T:\mathbb{B}(0,1)\longrightarrow\mathbb{H}$ be an operator which satisfies* (18), *$x^*\in B(0,1)$ and let $A:\mathbb{H}\longrightarrow\mathbb{H}$ be a bounded linear operator such that*

$$\forall h^*\in\mathbb{H},\quad \lim_{t\to 0}\frac{\|T(x^*+th^*)-T(x^*)-tAh^*\|}{t}=0,$$

then $$\forall h^*\in\mathbb{H},\quad \Delta_S'(T(x^*);Ah^*)=\langle x^*,Ah^*\rangle$$

and $$\partial\Delta_S(T(x^*))-\partial\Delta_S(T(x^*))\subset\{w^*\in\mathbb{H}\mid Aw^*=0\}.$$

Proof. In fact, let us fix $v^*\in\mathbb{H}\setminus\{0\}$. For $t\geq 0$ put $f(t):=\|v^*\|t$ and

$$g(t):=\begin{cases}1+\frac{\|T(x^*+tv^*)-T(x^*)\|+\|T(x^*-tv^*)-T(x^*)\|}{t}, & \text{if } \max\{\|x^*+tv^*\|,\|x^*-tv^*\|\}<1,\\ \infty, & \text{if } 1\leq\max\{\|x^*+tv^*\|,\|x^*-tv^*\|\}.\end{cases}$$

it follows from Corollary 5.1 that $\Delta_S'(T(x^*);Ah^*)=\langle x^*,Ah^*\rangle$ and $\langle x^*,A(-h^*)\rangle=\Delta_S'(T(x^*);-Ah^*)$. As a consequence $\langle y^*,Ah^*\rangle=\langle x^*,Ah^*\rangle=\Delta_S'(T(x^*);Ah^*)$ for all $y^*\in\partial\Delta_S(T(x^*))$, by Corollary 4.7. Thus the symmetry of A, see Fact 5.2, yields

$$\forall h^*\in\mathbb{H},\ \forall y^*\in\partial\Delta_S(T(x^*)),\ \forall z^*\in\partial\Delta_S(T(x^*)),\ \ 0=\langle y^*-z^*,Ah^*\rangle=\langle A(y^*-z^*),h^*\rangle.$$

Hence $\|A(y^*-z^*)\|=0$ for all $y^*,z^*\in\partial\Delta_S(T(x^*))$ and the inclusion

$$\partial\Delta_S(T(x^*))-\partial\Delta_S(T(x^*))\subset\{w^*\in\mathbb{H}\mid Aw^*=0\}$$

is established. $\square$

6. Directional convergence of subgradients

The Fréchet or Gâteaux differentiability of T allows us to say more on the strong convergence of some ε-subgradients of Δ_S, whenever $\varepsilon \downarrow 0$ or on the strong convergence of sugradients of Δ_S along some curves. A first observation is due to the fact that if a non-decreasing sequence of convex functions is pointwise converging to a function which is Fréchet differentiable at a given point, then subgradients of these functions strongly converge to the derivative, a version of this property is established in the proof of the following

Proposition 6.1. *Let $\mathbb{H}$ be a Hilbert space and $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18)*, $x^* \in \mathbb{B}(0,1)$ and $\{s_1, s_2, \ldots\} \subset S$ be such that*

$$\Delta_S(T(x^*)) = \lim_{i\to\infty} \|T(x^*) - s_i\|.$$

Assume that $A : \mathbb{H} \longrightarrow \mathbb{H}$ is a bounded linear operator such that $A(\mathbb{H})$ is a closed linear subspace of $\mathbb{H}$ and

$$\lim_{t\to 0} \sup_{h^*\in\mathbb{B}[0,1]} \frac{\|T(x^*+th^*) - T(x^*) - tAh^*\|}{t} = 0.$$

For each $i \in \mathbb{N}$ put $S_i := \{s_1, \ldots, s_i\}$.

Suppose that the sequences $\{x_i\}_{i\in\mathbb{N}}$ and $\{x_i^*\}_{i\in\mathbb{N}}$ satisfy $x_i^* \in \partial\Delta_{S_i}(x_i)$ for all $i \in \mathbb{N}$ and $\lim_{i\to\infty} \|x_i - T(x^*)\| = 0$. Then:

(i) $\lim_{i\to\infty} \|P_{A(\mathbb{H})}\Big(x_i^* - x^*\Big)\| = 0$;

(ii) *if* $\ker A$ *is finite dimensional and*

$$\forall i \in \mathbb{N}, \quad \Delta_{S_i}(T(x^*)) = \|T(x^*) - s_i\|,$$

then the sequence $\{s_i\}_{i\in\mathbb{N}}$ has a convergent subsequence;

(iii) *if* $\ker A$ *is finite dimensional and for $y^* \in \mathbb{H}$, $w_0 \in \mathbb{H}$*

$$\forall h \in \mathbb{S}_{\mathbb{H}}[0,1], \lim_{t\downarrow 0} \frac{\Delta'_S(x; w_0 + th) - \Delta'_S(x; w_0) - \langle y^*, th\rangle}{t} = 0,$$

and there are $t_i > 0$ for all $i \in \mathbb{N}$ such that $\lim_{i\to\infty} t_i = 0$*, and*

$$\forall i \in \mathbb{N}, \ \Delta_{S_i}(T(x^*)) = \|T(x^*) - s_i\|,$$

and (see Corollary 4.2 (iii))

$$\forall i \in \mathbb{N}, \ \frac{1}{\Delta_S(T(x^*) + t_i w_0)}(T(x^*) + t_i w_0 - s_i) \in \partial\Delta_S(T(x^*) + t_i w_0),$$

then the sequence $\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*) + t_i w_0 - s_i)\}_{i\in\mathbb{N}}$ converges to y^*.

Proof. (i) Observe that $\Delta_{S_i}(y) \le \Delta_{S_{i+1}}(y)$ for all $i \in \mathbb{N}$, $y \in \mathbb{H}$ and

$$\forall y \in \mathbb{H}, \lim_{i\to\infty} \Delta_{S_i}(y) \le \Delta_S(y) \text{ and } \lim_{i\to\infty} \Delta_{S_i}(x_i) = \Delta_S(T(x^*)).$$

Suppose that $\quad 0 < \lim_{i\to\infty} \|P_{A(\mathbb{H})}\Big(x_i^* - x^*\Big)\|.$

It follows from the Open Mapping Theorem, see [24, Theorem 1, p. 75], that there is $\delta > 0$ for which

$$\{y \in A(\mathbb{H}) \mid \|y\| < 2\delta\} \subset \{Ax \mid \|x\| < 1\}.$$

There are $\varepsilon >$ and $v_i^* \in \mathbb{B}[0,1]$ such that for all $i \in \mathbb{N}$ large enough we have

$$\varepsilon < \|x_i^* - x^*\| \text{ and } \varepsilon\delta < \langle x_i^* - x^*, Av_i^*\rangle.$$

For all $t > 0$ such that $\|x^* + tv_i^*\| < 1$ and $i \in \mathbb{N}$ large enough we have

$$\begin{aligned}
&\langle x_i^*, tAv_i^*\rangle \le \Delta_{S_i}(x_i + tAv_i^*) - \Delta_{S_i}(x_i)) \le \Delta_S(T(x^*) + tAv_i^*) - \Delta_S(T(x^*)) \\
&\quad + \|x_i - T(x^*)\| + \Delta_S(T(x^*)) - \Delta_{S_i}(x_i) \le \Delta_S(T(x^* + tv_i^*)) - \Delta_S(T(x^*)) \\
&\quad + \|x_i - T(x^*)\| + \Delta_S(T(x^*)) - \Delta_{S_i}(x_i) + \sup_{h^* \in \mathbb{B}[0,1]} \{\|T(x^*+th^*) - T(x^*) - tAh^*\|\} \\
&\le \langle x^* + tv_i^*, (T(x^* + tv_i^*)) - T(x^*))\rangle + \|x_i - T(x^*)\| + \Delta_S(T(x^*)) \\
&\quad - \Delta_{S_i}(x_i) + \sup_{h^* \in \mathbb{B}[0,1]} \{\|T(x^* + th^*) - T(x^*) - tAh^*\|\} \\
&\le \langle x^* + tv_i^*, tAv_i^*\rangle + \|x_i - T(x^*)\| + \Delta_S(T(x^*)) - \Delta_{S_i}(x_i) \\
&\quad + 2 \sup_{h^* \in \mathbb{B}[0,1]} \{\|T(x^* + th^*) - T(x^*) - tAh^*\|\},
\end{aligned}$$

hence

$$\begin{aligned}
\varepsilon\delta < \langle x_i^* - x^*, Av_i^*\rangle \le \frac{1}{t}\Big(&\langle tv_i^*, tAv_i^*\rangle + \|x_i - T(x^*)\| + \Delta_S(T(x^*)) - \Delta_{S_i}(x_i) \\
&+ 2 \sup_{h^* \in \mathbb{B}[0,1]} \{\|T(x^* + th^*) - T(x^*) - tAh^*\|\}\Big).
\end{aligned}$$

Thus we get the following contradiction

$$\begin{aligned}
0 < \varepsilon\delta \le \lim_{t \downarrow 0} \Big(&\limsup_{i \to \infty} \frac{1}{t}\Big(\langle tv_i^*, tAv_i^*\rangle + \|x_i - T(x^*)\| \\
&+ \Delta_S(T(x^*)) - \Delta_{S_i}(x_i) + 2 \sup_{h^* \in \mathbb{B}[0,1]} \{\|T(x^* + th^*) - T(x^*) - tAh^*\|\}\Big)\Big) = 0,
\end{aligned}$$

so $\lim_{i \to \infty} \|P_{A(\mathbb{H})}\Big(x_i^* - x^*\Big)\| = 0$.

(ii) If $\Delta_S(T(x^*)) = 0$, then $\lim_{i \to \infty} s_i = T(x^*)$ in this case. Let us consider the case $\Delta_S(T(x^*)) > 0$, observe that the assumption

$$\forall i \in \mathbb{N},\ \Delta_{S_i}(T(x^*)) = \|T(x^*) - s_i\|$$

yields

$$\forall i \in \mathbb{N}, \quad \|T(x^*) - s_i\| > 0 \Longrightarrow \|T(x^*) - s_i\|^{-1}(T(x^*) - s_i) \in \partial\Delta_{S_i}(T(x^*)).$$

Thus $\lim_{i \to \infty} \|P_{A(\mathbb{H})}\Big(\|T(x^*) - s_i\|^{-1}(T(x^*) - s_i) - x^*\Big)\| = 0$.

If $\ker A$ is finite dimensional, then there is an increasing sequence of integers $\{i_k\}_{k\in\mathbb{N}}$ such that the subsequence

$$\{P_{\{w^*\in\mathbb{H}|Aw^*=0\}}(\|T(x^*)-s_{i_k}\|^{-1}(T(x^*)-s_{i_k})-x^*)\}_{k\in\mathbb{N}}$$

is converging (with respect to the norm), which ensures the strong convergence of the subsequence $\{s_{i_k}\}$. This justifies the statement in (ii), see Fact 2.3.

(iii) It follows from Lemma 2.2 that

$$w-\lim_{t\downarrow 0}\mathbb{B}_{X^*}[0,M]\cap\partial\Delta_S(x+tw_0)=\{y^*\}. \tag{28}$$

Let $s_i\in S$, $t_i>0$ for all $i\in\mathbb{N}$ be such that $\lim_{i\to\infty}t_i=0$, and

$$\forall i\in\mathbb{N},\quad \Delta_{S_i}(T(x^*))=\|T(x^*)-s_i\|$$

and $$\frac{1}{\Delta_S(T(x^*)+t_iw_0)}\,(T(x^*)+t_iw_0-s_i)\in\partial\Delta_S(T(x^*)+t_iw_0).$$

It follows from the part of the proof concerning (ii) that any subsequence of the sequence $\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i)\}_{i\in\mathbb{N}}$ has a convergent (strongly) subsequence. It follows from (28) that its limit must be y^*, so the sequence

$$\left\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i)\right\}_{i\in\mathbb{N}}$$

converges to y^*. □

Instead of the Fréchet differentiability of T the Gâteaux differentiability can be assumed to get similar convergence results.

Proposition 6.2. *Let $\mathbb{H}$ be a Hilbert space and $T:\mathbb{B}(0,1)\longrightarrow\mathbb{H}$ be an operator which satisfies* (18)*, $x^*\in\mathbb{B}(0,1)$ and $\{s_1,s_2,\ldots\}\subset S$, $s_0\in S$ and $\Delta_S(T(x^*))=\|T(X^*)-s_0\|$. Assume that $A:\mathbb{H}\longrightarrow\mathbb{H}$ is a bounded linear operator such that $A(\mathbb{H})$ is a closed linear subspace of $\mathbb{H}$ and* $\ker A$ *is finite dimensional, and*

$$\forall h^*\in\mathbb{H},\ \lim_{t\to 0}\frac{\|T(x^*+th^*)-T(x^*)-tAh^*\|}{t}=0.$$

For each $i\in\mathbb{N}$ put $S_i:=\{s_0,s_1,\ldots,s_i\}$.

Let $w_0\in\mathbb{H}$ be given $t_i>0$, $\varepsilon_i>0$ for all $i\in\mathbb{N}$ be such that $\lim_{i\to\infty}t_i=0$, $\lim_{i\to\infty}\varepsilon_i=0$, and

$$\forall i\in\mathbb{N},\ \forall v\in\mathbb{H},\ \Delta'_{S_i}(T(x^*)+t_iw_0;v)\le\varepsilon_i\|v\|+\Delta'_{S_i}(T(x^*);v). \tag{29}$$

The following statements hold true:

(i) *if $x_i^*\in\partial\Delta_{S_i}(T(x^*)+t_iw_0)$ for all $i\in\mathbb{N}$, then*

$$\lim_{i\to\infty}\|P_{A(\mathbb{H})}\Big(x_i^*-x^*\Big)\|=0$$

and the sequence $\{x_i^*\}_{i\in\mathbb{N}}$ has a convergent subsequence;

(ii) *if* $\forall i \in \mathbb{N}$, $\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i) \in \partial\Delta_S(T(x^*)+t_iw_0)$, *and for* $y^* \in \mathbb{H}$

$$\forall h \in \mathbb{S}_X, \quad \lim_{t\downarrow 0} \frac{\Delta'_S(x;w_0+th)-\Delta'_S(x;w_0)-\langle y^*,th\rangle}{t} = 0,$$

then the sequence $\left\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i)\right\}_{i\in\mathbb{N}}$ *converges to* y^*.

Proof. (i) Observe that by (29) we get $\forall i \in \mathbb{N}$, $\forall v \in \mathbb{H}$,

$$\Delta'_{S_i}(T(x^*)+t_iw_0;v) \le \varepsilon_i\|v\| + \Delta'_{S_i}(T(x^*);v) \le \varepsilon_i\|v\| + \Delta'_S(T(x^*);v).$$

If $x_i^* \in \partial\Delta_{S_i}(T(x^*)+t_iw_0)$ for all $i \in \mathbb{N}$, then by Corollary 4.7 we get

$$\begin{aligned}\forall i \in \mathbb{N},\ \forall h^* \in \mathbb{H},\ \langle x_i^*, Ah^*\rangle &\le \Delta'_{S_i}(T(x^*)+t_iw_0;Ah^*)\\ &\le \varepsilon_i\|Ah^*\| + \Delta'_S(T(x^*);Ah^*) = \varepsilon_i\|Ah^*\| + \langle x^*, Ah^*\rangle.\end{aligned}$$

Thus $$0 = \lim_{i\to\infty}\left(\sup_{h^*\in\mathbb{B}[0,1]} \langle A(x_i^*-x^*),h^*\rangle\right) = \lim_{i\to\infty}\|A(x_i^*-x^*)\|,$$

which together with Fact 2.3 yields $\lim_{i\to\infty}\|P_{A(\mathbb{H})}\left(x_i^*-x^*\right)\| = 0$.

The possibility of the choice of the convergent subsequence is due to the fact that $\mathbb{H} = A(H) + \ker A$, AH is closed and $\ker A$ is finite dimensional.

(ii) It follows from Lemma 2.2 that

$$w-\lim_{t\downarrow 0}\mathbb{B}_{X^*}[0,M] \cap \partial\Delta_S(x+tw_0) = \{y^*\}. \tag{30}$$

It follows from the part of the proof concerning (i) that any subsequence of the sequence $\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i)\}_{i\in\mathbb{N}}$ has a convergent (strongly) subsequence. It follows from (30) that its limit must be y^*, so the sequence

$$\left\{\frac{1}{\Delta_S(T(x^*)+t_iw_0)}(T(x^*)+t_iw_0-s_i)\right\}_{i\in\mathbb{N}}$$

converges to y^*. □

Let $Z \subset \mathbb{H}$ be a compact subset and $x, w_0 \in \mathbb{H}$ be given with $\Delta_Z(x) > 0$ and $t_i > 0$, $\varepsilon_i > 0$ for all $i \in \mathbb{N}$ be such that $\lim_{i\to\infty} t_i = 0$. Take $z_i \in Z$ such that $\|x+t_iw_0-z_i\| = \Delta_Z(x+t_iw_0)$ for all $i \in \mathbb{N}$. It is easy to verify that

$$\limsup_{i\to\infty}\left(\sup_{v\in\mathbb{B}[0,1]} \frac{\langle x+t_iw_0-z_i,v\rangle}{\Delta_Z(x)} - \Delta'_Z(x;v)\right) \le 0,$$

thus assuming that w_0 and $t_1, t_2, \dots$ are such that $D_F\Delta_Z(x+t_iw_0)$ exists for each $i \in \mathbb{N}$, see Corollary 4.2(iii), we get

$$\limsup_{i\to\infty}\left(\sup_{v\in\mathbb{B}[0,1]} \Delta'_Z(x+t_iw_0;v) - \Delta'_Z(x;v)\right) \le 0.$$

this means that the inequalities in (29) hold true for a residual set of directions. Finally a result preserving the strong convergence of subgradients along some curves, whenever the Fréchet derivative of T exists at x^* and has a finite dimensional kernel, is given, see Corollary 6.4. Namely, under the assumption that the kernel of the second order derivative of Δ_S^* at x^* is a finite dimensional space, we get

$$\forall h^* \in \mathbb{H}, \limsup_{t\downarrow 0} \partial\Delta_S(T(x^*+th^*)) \neq \emptyset.$$

Proposition 6.3. *Let $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18), *let $x^* \in \mathbb{B}(0,1)$ and $A : \mathbb{H} \longrightarrow \mathbb{H}$ be a bounded linear operator such that*

$$\lim_{t\to 0} \sup_{h^*\in\mathbb{B}[0,1]} \frac{\|T(x^*+th^*) - T(x^*) - tAh^*\|}{t} = 0, \tag{31}$$

then
$$\forall h^* \in \mathbb{H}, \limsup_{t\downarrow 0} \left(\sup_{y^*\in\partial\Delta_S(T(x^*+th^*))} \|A(y^*-x^*)\| = 0 \right). \tag{32}$$

Moreover, if $A(\mathbb{H})$ is a closed linear subspace of $\mathbb{H}$, then

$$\forall h^* \in \mathbb{H}, \limsup_{t\downarrow 0} \left(\sup_{y^*\in\partial\Delta_S(T(x^*+th^*))} d_{\{w^*\in\mathbb{H}|Aw^*=0\}}(y^*-x^*) = 0 \right).$$

Proof. Suppose that (32) is not valid, that is there are $\varepsilon > 0$, $h^* \in \mathbb{H} \setminus \{0\}$ and sequences $\{t_i\}_{i\in\mathbb{N}}$, $\{y_i^*\}_{i\in\mathbb{N}}$,$\{w_i\}_{i\in\mathbb{N}}$ such that $t_i > 0$ for all $i \in \mathbb{N}$, $\lim_{i\to\infty} t_i = 0$ and

$$\forall i \in \mathbb{N}, \quad y_i^* \in \partial\Delta_S(T(x^*+t_ih^*))$$

and
$$2\varepsilon < \|A(y_i^*-x^*)\| = \langle A(y_i^*-x^*), w_i\rangle, \|w_i\| = 1.$$

We may assume that

$$\|T(x^*+t_i(h^*+w_i)) - T(x^*) - t_iA(h_i^*+w_i)\| \leq \frac{3\varepsilon t_i}{8}$$

and
$$\|T(x^*+t_ih^*) - T(x^*) - t_iAh_i^*\| \leq \frac{3\varepsilon t_i}{8}$$

for all $i \in \mathbb{N}$. Thus for all $i \in \mathbb{N}$ large enough it follows that

$$\begin{aligned}
t_i(2\varepsilon+\langle x^*, Aw_i\rangle) \leq t_i\langle y_i^*, Aw_i\rangle &\leq \frac{3\varepsilon t_i}{4} + \langle y_i^*, T(x^*+t_i(h^*+w_i)) - T(x^*+t_ih^*)\rangle \\
&\leq \frac{3\varepsilon t_i}{4} + \Delta_S(T(x^*+t_i(h^*+w_i))) - \Delta_S(T(x^*+t_ih^*))\rangle \\
&\leq \frac{3\varepsilon t_i}{4} + \langle x^*+t_i(h^*+w_i), T(x^*+t_i(h^*+w_i)) - T(x^*+t_ih^*)\rangle \\
&\leq t_i\Big(\varepsilon + \langle x^*+t_i(h^*+w_i), Aw_i\rangle\Big),
\end{aligned}$$

hence we get the following contradiction

$$\varepsilon \leq \langle t_i(h^*+w_i^*), Aw_i\rangle$$

whenever $i \in \mathbb{N}$ is large enough. Thus (32) holds true. The second part of the statement is a simple consequence of the first one and Fact 2.3. □

Corollary 6.4. *Let $T : \mathbb{B}(0,1) \longrightarrow \mathbb{H}$ be an operator which satisfies* (18)*, let $x^* \in \mathbb{B}(0,1)$ and $A : \mathbb{H} \longrightarrow \mathbb{H}$ be a bounded linear operator such that $A(\mathbb{H}) = \operatorname{cl} A(\mathbb{H})$, and* $\ker A$ *is finite dimensional, and*

$$\lim_{t\to 0} \sup_{h\in\mathbb{B}[0,1]} \frac{\|T(x^*+th^*) - T(x^*) - tAh^*\|}{t} = 0.$$

Then for every $h^* \in \mathbb{H} \setminus \{0\}$ and sequences $\{t_i\}_{i\in\mathbb{N}}$, $\{y_i^*\}_{i\in\mathbb{N}}$ such that $t_i > 0$ for all $i \in \mathbb{N}$, $\lim_{i\to\infty} t_i = 0$ and

$$\forall i \in \mathbb{N}, y_i^* \in \partial\Delta_S(T(x^*+t_ih^*)),$$

there are $y^* \in \partial\Delta_S(T(x^*))$ and an increasing sequence of integers $\{i_k\}_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \|y_{i_k}^* - y^*\| = 0.$$

Proof. Let us fix $h^* \in \mathbb{H} \setminus \{0\}$ and sequences $\{t_i\}_{i\in\mathbb{N}}$, $\{y_i^*\}_{i\in\mathbb{N}}$ such that

$$\forall i \in \mathbb{N}, \quad y_i^* \in \partial\Delta_S(T(x^*+t_ih^*)).$$

It follows from Proposition 6.3 that $\lim_{i\to\infty} d_{\{w^*\in\mathbb{H}\mid Aw^*=0\}}(y_i^* - x^*) = 0.$

Thus $\lim_{i\to\infty} \|P_{A(\mathbb{H})}(y_i^* - x^*)\| = 0$, see Fact 2.3.

Since the space $\{w^* \in \mathbb{H} \mid Aw^* = 0\}$ is finite dimensional, so there is an increasing sequence of integers $\{i_k\}_{k\in\mathbb{N}}$ such that the subsequence $\{P_{\{w^*\in\mathbb{H}\mid Aw^*=0\}}(y_{i_k}^* - x^*)\}_{k\in\mathbb{N}}$ is converging (with respect to the norm). This justifies the statement. □

References

[1] H. H. Bauschke, P. L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, Berlin (2011).

[2] J. M. Borwein: *Future challenges for variational analysis, variational analysis and generalized differentiation in optimization and control*, Optimization Appl. 47 (2010) 95–105.

[3] J. M. Borwein, D. Noll: *Second order differentiability of convex functions in Banach spaces*, Trans. Amer. Math. Soc. 342 (1994) 43–81.

[4] J. M. Borwein, J. D. Vanderwerff: *Convex Functions: Constructions, Characterizations and Counterexamples*, Cambridge University Press, Cambridge (2010).

[5] J. M. Borwein, Q. J. Zhu: *A survey of subdifferential calculus with applications*, Nonlinear Analysis 38 (1999) 687–773.

[6] S. Fitzpatrick: *Metric projections and the differentiability of distance functions*, Bull. Austral. Math. Soc. 22 (1980) 291–312.

[7] F. Giannessi: *A problem on convex functions*, J. Optim. Theory Appl. 59 (1988) 525.

[8] D. H. Griffel: *Applied Functional Analysis*, Ellis Horwood, Hemel (1985).

[9] J.-B. Hiriart-Urruty: *Potpourri of conjectures and open questions in nonlinear analysis and optimization*, SIAM Review 49 (2007) 255–273.

[10] A. Jourani, L. Thibault, D. Zagrodny: *$C^{1,\omega(\cdot)}$-regularity and Lipschitz-like properties of subdifferential*, Proc. London Math. Soc. 105 (2012) 189–223.

[11] A. Jourani, L. Thibault, D. Zagrodny: *Differential properties of the Moreau envelope*, J. Funct. Analysis 266 (2014) 1185–1237.

[12] A. Jourani, L. Thibault, D. Zagrodny: *The NSLUC property and Klee envelope*, Math. Ann. 365/3 (2016) 923–967.

[13] J. Lindenstrauss, D. Preiss, J. Tišer: *Fréchet Differentiability of Lipschitz Functions and Porous Sets in Banach Space*, Princeton University Press, Princeton (2012).

[14] F. Mignot: *Contrôle dans les inéquations variationelles elliptiques*, J. Funct. Analysis 22 (1976) 130–185.

[15] M. Moussaoui, A. Seeger: *Second-order subgradients of convex integral functional*, Trans. Amer. Math. Soc. 351 (1999) 3687–3711.

[16] J. C. Oxtoby: *Measure and Category. A Survey of the Analogies between Topological and Measure Spaces*, 2nd ed., Springer, New York (1980).

[17] R. R. Phelps: *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Springer, Berlin (1993).

[18] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, Springer, Berlin (1997).

[19] R. T. Rockafellar,D. Zagrodny: *A derivative-coderivative inclusion in second-order nonsmooth analysis*, Set-Valued Analysis 5 (1997) 89–105.

[20] W. Rudin: *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York (1976).

[21] W. Rudin: *Functional Analysis*, 2nd ed., McGraw-Hill, New York (1991).

[22] S. Simons: *Minimax and Monotonicity*, Lecture Notes in Mathematics 1693, Springer, Berlin (1998).

[23] U. Westphal, T. Schwartz: *Farthest points and monotone operators*, Bull. Austral. Math. Soc. 58 (1998) 75–92.

[24] K. Yosida: *Functional Analysis*, 6th ed., Springer, Berlin (1995).

[25] D. Zagrodny: *On the strong convergence of subgradients of convex functions*, J. Optim. Theory Appl. 178 (2018) 441–423.

[26] D. Zagrodny: *The strong convergence of subgradients of convex functions along directions: Perspectives and open problems*, J. Optim. Theory Appl. 178 (2018) 660–671.