Journal of Convex Analysis
Volume 11 (2004), No. 2, 335–361

Variational Analysis for a Class of Minimal Time Functions in Hilbert Spaces

Giovanni Colombo*
Dipartimento di Matematica Pura e Applicata,
Università di Padova, via Belzoni 7, 35131 Padova, Italy
colombo@math.unipd.it

Peter R. Wolenski†
Department of Mathematics, Louisiana State University,
326 Lockett Hall, Baton Rouge, Louisiana 70803-4918, USA
wolenski@math.lsu.edu

Received May 20, 2003
Revised manuscript received January 8, 2004

This paper considers the parameterized infinite dimensional optimization problem

$$\text{minimize} \quad \big\{t \geq 0 : \; S \cap \{x + tF\} \neq \emptyset\big\},$$

where S is a nonempty closed subset of a Hilbert space H and $F \subseteq H$ is closed convex satisfying $0 \in \operatorname{int} F$. The optimal value $T(x)$ depends on the parameter $x \in H$, and the (possibly empty) set $S \cap (x + T(x)F)$ of optimal solutions is the "F-projection" of x into S. We first compute proximal and Fréchet subgradients of $T(\cdot)$ in terms of normal vectors to level sets, and secondly, in terms of the F-projection. Sufficient conditions are also obtained for the differentiability and semiconvexity of $T(\cdot)$, results which extend the known case when F is the unit ball.

1. Introduction

Let H be a real Hilbert space, and suppose the following data are given: $S \subseteq H$ is nonempty and closed, and $F \subseteq H$ is nonempty, closed, convex, bounded, and with $0 \in \operatorname{int} F$. This paper studies regularity properties of the value function $T : H \to I\!R$ given by

$$T(x) := \inf_{t \geq 0} \big\{t : S \cap \{x + tF\} \neq \emptyset\big\}. \tag{1}$$

One can view $T(\cdot)$ as the *minimal time function* of a control system in which the dynamic equation is $\dot{x}(t) \in F$, where the righthand side F is constant (i.e. independent of t and x), and the target set is S. An important and well-studied special case is $F = \overline{I\!B}$, where $\overline{I\!B}$ is the closed unit ball, in which $T(\cdot)$ reduces to the usual distance function

$$d_S(x) = \inf_{s \in S} \|x - s\|.$$

We adopt the approach of Cannarsa and Sinestrari [8] by studying $T(\cdot)$ in analogy with $d_S(\cdot)$. A key ingredient to differentiability properties of $d_S(\cdot)$ is the set of projections onto

*work supported by Italian MURST (COFIN project "Analysis and control of deterministic and stochastic evolution equations") and GNAMPA (project "Partial differential equations and control theory")
†work supported in part by NSF grant DMS-9972241

ISSN 0944-6532 / $ 2.50

S; the analogous notion we require is the set $\Pi_S^F(x)$ of so-called F-projections onto S, given by $\Pi_S^F(x) = S \cap \{x + T(x)F\}$, which could be empty in infinite dimensions.

We quickly review some recent work on minimal time problems related to our results here. The minimal time function for general nonlinear control problems was first characterized as a solution to a Hamilton-Jacobi equation by Bardi [3] using viscosity methods. Soravia [18] extended these results to allow for noncontrollability and more general boundary conditions; see also [21] for related results using an invariance-based approach. Regularity properties were derived in [8] with an emphasis on *global* semiconcavity results. Global semiconvexity was also proven in [8] for the case where the target is convex and the dynamic equation is linear. Lipschitz estimates are provided by Veliov [20], and differentiability is linked with properties of the velocity set in [5], where the dynamics are linear and the target is a single point. The underlying state space in the aforementioned papers is finite-dimensional. Some infinite dimensional extensions of [8] are provided in [9], where in particular sufficient conditions are established for differentiability. The results of Section 5.3 below are contained in [9] when $S = \{0\}$. Other regularity results, for a more general dynamics but less general target and control set, are contained in [1]. As already mentioned, the case $F = \overline{I\!\!B}$ has $T(\cdot) = d_S(\cdot)$, and has been well-studied in both finite and infinite dimensions. Differentiability in Hilbert space of $d_S(\cdot)$ is *characterized* in [12]. Additional refinements and localization is carried out in [16], and [13] contains related results emphasizing the distinction between finite and infinite dimensional spaces.

The present paper extends to $T(\cdot)$ a variety of the above-mentioned results. The literature on $d_S(\cdot)$, both in finite and in infinite dimensions, goes much deeper (see, e.g., [7] and references therein) than we are able to establish for $T(\cdot)$. However, we compute its subdifferentials, and derive regularity properties under suitable assumptions on both S and F. The linkage of regularity of $T(\cdot)$ with the interplay of the regularity properties of S and F at the point of contact is firmly established.

We remark that our class of problems cannot be simply formulated as a renorming of the space, unless F is symmetric with respect to the origin. A source of difficulty lies in the fact that the segment joining x with its F-projection is not necessarily normal to S or to F.

The plan of the paper is as follows. Section 2 contains a terse review of the required background plus some preliminary results. The main result in Section 3 is a general formula for the proximal and the Fréchet subgradient of $T(\cdot)$ in terms of normal vectors to its level sets. A similar propagation formula for general nonlinear systems in finite dimensions was proven by Soravia [19]; proximal versions in both finite and infinite dimensions are contained in [21] and [14]. Similar results are also contained in §3 of [4]. Section 4 contains a formula for the subgradient of $T(\cdot)$ in the case where S is convex. Section 5 contains the main results, and considers nonconvex target sets. If the target S is not weakly closed, then the nonemptiness of the F-projection is not assured, and so we first provide sufficient conditions for existence, uniqueness, and local Lipschitz continuity of the F-projection (5.1); these generalize well known assertions for the Euclidean metric projection (see [6, 12, 16, 13]). Again analogous to the distance function (see e.g. [12, 16]), results on the differentiability of $T(\cdot)$ are derived there as well. In particular, we show in Theorem 5.14 that if F is strictly convex and $\mathcal{C}^{1,+}$, and S is proximally smooth with the nonconvexity of S "harmonious" with the strict convexity of F, then $T(\cdot)$ is of class $\mathcal{C}^{1,+}$ in a neighborhood of S. Moreover, assumptions are introduced that imply semiconvexity in a neighborhood of S. Finally, Section 6 is devoted to some examples and remarks.

2. Preliminaries

This section reviews some of the concepts and basic tools used in the sequel. See [11] for a fuller development of nonsmooth analysis based on the proximal concepts, and [17] for a somewhat different but exhaustive treatment in finite dimensions.

2.1. Background in variational analysis

The *proximal normal cone* $N_S^p(s)$ to S at $s \in S$ is the set of all $\zeta \in H$ for which there exist $\sigma \geq 0$ such that

$$\langle \zeta, s' - s \rangle \leq \sigma \|s' - s\|^2 \quad \forall s' \in S.$$

If S is convex, then the proximal normal cone coincides with the normal cone $N_S(s)$ of convex analysis, and in this case, there is no loss in generality using just $\sigma = 0$.

The corresponding function concept is the proximal subgradient, defined as follows. Suppose $f : H \to (-\infty, \infty]$ is lower semicontinuous and proper, and let $\operatorname{epi} f := \{(x, \alpha) \in H \times I\!R : \alpha \geq f(x)\}$ denote the epigraph of f. For $x \in \operatorname{dom} f := \{x \in H : f(x) < \infty\}$, the proximal subgradient $\partial_p f(x)$ is (the possibly empty) subset of H defined as those ζ satisfying $(\zeta, -1) \in N^P_{\operatorname{epi} f}\big(x, f(x)\big)$. A user-friendly description of the proximal subgradient is given by (see [11, Theorem 2.5])

$$\partial_p f(x) = \{\zeta : \exists \eta > 0,\ \sigma \geq 0 \text{ so that } f(y) \geq f(x) + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \ \ \forall y \in x + \eta \overline{I\!B}\}.$$

If f is convex, then $\partial_p f(x)$ coincides with the subgradient of convex analysis, and is simply denoted by ∂f. In this case, the above description is equivalent with $\sigma = 0$ and $\eta = \infty$.

Additional variational objects will also be considered. The *Fréchet normal cone* $N_S^f(s)$ to S at s is the set of all $\zeta \in H$ such that

$$\limsup_{s' \to s, s' \in S} \left\langle \zeta, \frac{s' - s}{\|s' - s\|} \right\rangle \leq 0.$$

The corresponding function concept $\partial_f f(x)$ can be defined via the epigraph as in the proximal case, or equivalently by

$$\zeta \in \partial_f f(x) \qquad \text{if and only if} \qquad \liminf_{x' \to x} \frac{f(x') - f(x) - \langle \zeta, x' - x \rangle}{\|x' - x\|} \geq 0.$$

Limiting objects are also useful: the *limiting normal cone* to S at s is the set N_S^l of all cluster points in the weak topology of H of sequences $\{\zeta_i\}$ with $\zeta_i \in N_S^p(s_i)$ and $s_i \to s$, i.e., $N_S^l(s) = \{\zeta : \zeta = w - \lim \zeta_i,\ \zeta_i \in N_S^p(s_i),\ s_i \to s\}$. The limiting subdifferential of a lower semicontinuous proper function $f : H \to (-\infty, +\infty]$ at $x \in \operatorname{dom} f$ is the set $\partial_l f(x)$ of all weak cluster points of sequences $\{\zeta_i\}$ with $\zeta_i \in \partial_p f(x_i)$ and $x_i \to x$. The *Clarke normal cone* $N_S^c(s)$ is defined as the closed convex hull of $N_S^l(s)$. It can be proved that $N_S^p(s) \subseteq N_S^f(s) \subseteq N_S^l(s) \subseteq N_S^c(s)$ for all $s \in S$, where all inclusions can be strict. Similar chains of inclusions hold also for subgradients (see, e.g, [11, §10], [2, Theorem 4.4.4]). Equalities among these normal cones/subgradients are entwined with regularity properties of the sets/functions, and are a major theme of the ensuing analysis.

2.2. Gauge functions and polars

We are interested in convex sets $F \subset H$ that are closed, convex, bounded, and with $0 \in \operatorname{int} F$, however the concepts of this section are first introduced more generally with F assumed only to be closed, convex, and with $0 \in F$.

Recall that the (Minkowski) *gauge* function $\rho_F : H \to [0, \infty]$ associated to F is defined by

$$\rho_F(\zeta) = \min\left\{ t \geq 0 : \frac{1}{t}\,\zeta \in F \right\},$$

and the polar F° of F is the set

$$F^\circ := \{\zeta \ : \ \langle \zeta, v\rangle \leq 1 \ \forall v \in F\}.$$

The polar is always closed, convex, and with $0 \in F^\circ$, and the closedness and convexity of F imply the polar of F° is F, that is $(F^\circ)^\circ = F$. We next further review some elementary properties of $\rho_F(\cdot)$, which of course also hold for $\rho_{F^\circ}(\cdot)$, but are not explicitly stated. The proofs of these facts involve routine manipulations of the definitions and are therefore omitted. The gauge $\rho_F(\cdot)$ is positively homogeneous ($\rho_F(rx) = r\rho_F(x)$ for all $x \in H$ and $r \geq 0$) and subadditive ($\rho_F(x + y) \leq \rho_F(x) + \rho_F(y)$ for all x and y), and therefore is also convex. Since F is closed, $\rho_F(\cdot)$ is lower semicontinuous. It is clear that $x \in F$ if and only if $\rho_F(x) \leq 1$. Furthermore, F is bounded if and only if $0 \in \operatorname{int} F^\circ$, and in fact one has for constants m and M in $[0, \infty]$ that

$$m\overline{\mathbb{B}} \subseteq F \subseteq M\overline{\mathbb{B}} \quad \iff \quad \frac{1}{M}\overline{\mathbb{B}} \subseteq F^\circ \subseteq \frac{1}{m}\overline{\mathbb{B}}. \tag{2}$$

Henceforth for the remainder of the paper, **we assume $\boldsymbol{F}$ is bounded and $\mathbf{0} \in \operatorname{\mathbf{int}} \boldsymbol{F}$**. Of course F bounded (resp. $0 \in \operatorname{int} F$) means that we can take $M < \infty$ (resp. $m > 0$) in (2). Note then that (2) implies F° is also bounded with $0 \in \operatorname{int} F^\circ$. Define, for a set $A \subseteq H$, $\|A\| = \sup\{\|a\| \ : \ a \in A\}$. Then, by (2) the number $\|F\|\,\|F^\circ\|$ is ≥ 1, and represents a measure of "how far" is F from being a ball. Some further consequences of these assumptions are included in the following proposition.

Proposition 2.1. *Suppose F is closed, convex, bounded, and with $0 \in \operatorname{int} F$. Then*

(*a*) $v \in \operatorname{bdry} F$ *if and only if* $\rho_F(v) = 1$.

(*b*) *For all* $\zeta \neq 0$ *in* H,
$$0 < \rho_{F^\circ}(\zeta) = \max_{v \in F} \langle \zeta, v\rangle < \infty.$$

(*c*) *Define* $\|F\| := \max\{\|v\| : v \in F\}$, *and similarly* $\|F^\circ\| := \max\{\|\zeta\| : \zeta \in F^\circ\}$. *Then for all* $z \in H$,
$$\frac{\rho_F(z)}{\|F^\circ\|} \leq \|z\| \leq \|F\|\rho_F(z). \tag{3}$$

Proof. (a) The "if" direction holds for any convex F with $0 \in F$, since $\rho_F(v) = 1$ implies $(1 + \varepsilon)v \notin F$ for all $\varepsilon > 0$. For the "only if" direction, we prove the contrapositive, and assume $\rho_F(v) \leq \rho_0 < 1$. Let $\varepsilon < \frac{1-\rho_0}{\max\{\rho_F(b') : b' \in \overline{\mathbb{B}}\}}$. Then for all $b \in \overline{\mathbb{B}}$, we have

$$\rho_F(v + \varepsilon b) \leq \rho_F(v) + \varepsilon\rho_F(b) \leq \rho_0 + \frac{1 - \rho_0}{\max\{\rho_F(b') : b' \in \overline{\mathbb{B}}\}}\rho_F(b) < 1.$$

Therefore $v + \varepsilon\overline{I\!B} \subseteq F$, and so $v \notin$ bdry F, and (a) is proven.

(b) Let $\zeta \neq 0$. Now $\rho_{F^\circ}(\zeta)$ is positive since $0 \in$ int F°, and is finite since F is bounded. A calculation shows

$$\rho_{F^\circ}(\zeta) = \min\left\{t : \frac{1}{t}\zeta \in F^\circ\right\} = \min\left\{t : \langle \zeta, v\rangle \leq t \ \forall v \in F\right\} = \max_{v \in F}\langle \zeta, v\rangle,$$

which proves (b).

(c) For any $z \in H$, we have by (b) that

$$\rho_F(z) = \max_{\zeta \in F^\circ}\langle \zeta, z\rangle \leq \|F^\circ\|\|z\|,$$

which is equivalent to the first inequality. The second is trivial if $z = 0$ and follows otherwise since $\frac{z}{\rho_F(z)} \in F$. □

Remark. Note (b) says that the Hamiltonian as usually defined in optimal control (see [11]) is $\rho_{F^\circ}(\cdot)$. Moreover, observe that ρ_F is a norm on H, equivalent to the usual one, if and only if $F = -F$. Finally, by (2) it is easy to see that $\|F\|\,\|F^\circ\| \geq 1$.

2.3. Subgradient properties of the gauge functions

Differential properties of the gauge functions are reviewed in this subsection, and in particular the duality relationships among them.

Proposition 2.2. *Assume that F is closed, convex, bounded, and with $0 \in$ int F. For given nonzero vectors $v \in H$, $\zeta \in H$, the following statements are equivalent:*

(a) $\langle \zeta, v\rangle = \rho_F(v)\rho_{F^\circ}(\zeta)$;

(b) $\frac{v}{\rho_F(v)}$ *attains the max over $u \in F$ of the map $u \mapsto \langle \zeta, u\rangle$;*

(c) $\zeta \in N_F\left(\frac{v}{\rho_F(v)}\right)$;

(d) $\frac{\zeta}{\rho_{F^\circ}(\zeta)}$ *attains the max over $\xi \in F^\circ$ of the map $\xi \mapsto \langle \xi, v\rangle$;*

(e) $v \in N_{F^\circ}\left(\frac{\zeta}{\rho_{F^\circ}(\zeta)}\right)$;

(f) $\frac{\zeta}{\rho_{F^\circ}(\zeta)} \in \partial\rho_F(v)$;

(g) $\frac{v}{\rho_F(v)} \in \partial\rho_{F^\circ}(\zeta)$.

Proof. By Proposition 2.1(b), one observes that $\rho_{F^\circ}(\cdot)$ is the support function of F, and so its Legendre-Fenchel conjugate

$$\rho_{F^\circ}(v) := \sup_{\zeta \in H}\left\{\langle \zeta, v\rangle - \rho_{F^\circ}(\zeta)\right\}$$

is the indicator function

$$I_F(v) := \begin{cases} 0 & \text{if } v \in F \\ \infty & \text{if } v \notin F \end{cases}$$

of F. Recall that the convex subgradient of $\partial I_F(v)$ is the normal cone to F at v. The equivalent conditions in the statement are immediate consequences of this fact and the duality relationships between a convex function and its conjugate. See e.g. [15, pp. 21-22]. □

Corollary 2.3. *Suppose $v \in H$. Then*

$$\partial \rho_F(v) = \left\{ \zeta : \rho_{F^\circ}(\zeta) = 1 \right\} \cap N_F \left(\frac{v}{\rho_F(v)} \right).$$

2.4. The minimal time function

The minimal time function $T(\cdot) : H \to [0, \infty)$ was defined above in (1), but the equivalent following description is more useful:

$$T(x) = \inf_{s \in S} \rho_F(s - x). \tag{4}$$

Note that the assumption $0 \in \operatorname{int} F$ implies $T(x) < \infty$ for all $x \in H$, but the next proposition says in fact $T(\cdot)$ is Lipschitz, as it is to be expected from the optimal control point of view (see [8], [20], [21]) since $0 \in \operatorname{int} F$.

Proposition 2.4. *The minimal time function $T(\cdot)$ is globally Lipschitz on H of rank $\|F^\circ\|$.*

Proof. Let $x, y \in H$ and let $\varepsilon > 0$. There exists $s \in S$ so that $\rho_F(s - y) \leq T(y) + \varepsilon$. By (4) and the subadditivity of $\rho_F(\cdot)$, we have

$$\begin{aligned} T(x) - T(y) &\leq \rho_F(s - x) - \rho_F(s - y) + \varepsilon \\ &\leq \rho_F(y - x) + \varepsilon \\ &\leq \|F^\circ\| \|y - x\| + \varepsilon, \end{aligned}$$

where in the last inequality we used (3). Letting $\varepsilon \downarrow 0$ and switching the roles of x and y proves the theorem. □

Another consequence of (3) is the following

Proposition 2.5. *Let $x \notin S$. Then $T(x) \leq \|F^\circ\| d_{bdry\, S}(x)$.*

The level sets $S(r)$ of $T(\cdot)$ will play a significant role in our analysis, and are defined by

$$S(r) = \{ y \in H : T(y) \leq r \}. \tag{5}$$

The F-projection of a point $x \in H$ is the (possibly empty) set $\Pi^F_S(x) = \{s \in S : \rho_F(s - x) = T(x)\}$. The metric projection onto a set A, i.e., the F-projection onto A when F is the closed unit ball, will be denoted by π_A. It is clear that the nonemptiness of $\Pi^F_S(x)$ means that the infimum in (1) is attained, i.e., there exists an optimal trajectory for the control problem (1). The following proposition contains the special versions of the so-called principle of optimality that are pertinent here.

Proposition 2.6 (Principle of Optimality). *Suppose $x \notin S$.*

(*a*) *For all $v \in F$ and $t \geq 0$,*

$$T(x - tv) \leq T(x) + t.$$

(*b*) *Let $S(r)$ be as in (5). Then*

$$T(x) \leq r + \min_{z \in S(r)} \rho_F(z - x).$$

Proof. (a). Let $v \in F$, $t \geq 0$, and $\varepsilon > 0$. There exists $s \in S$ so that $\rho_F(s-x) < T(x)+\varepsilon$. By subadditivity and positive homogeneity, we have

$$T(x - tv) \leq \rho_F(s - x + tv) \leq \rho_F(s - x) + t\rho_F(v) < T(x) + t + \varepsilon.$$

Letting $\varepsilon \downarrow 0$ proves (a).

(b). Let $\varepsilon > 0$ and suppose $0 \leq r < T(x)$, otherwise there is nothing to prove. There exist $z \in S(r)$ and $s \in S$ so that

$$\rho_F(z - x) < \min_{z' \in S(r)} \rho_F(z' - x) + \varepsilon \quad \text{and} \quad \rho_F(s - z) \leq r + \varepsilon.$$

Therefore

$$T(x) \leq \rho_F(s - z) + \rho_F(z - x) \leq r + \min_{z' \in S(r)} \rho_F(z' - x) + 2\varepsilon,$$

and letting $\varepsilon \downarrow 0$ proves (b). □

Remark. If $r \leq T(x)$, then the inequality in the statement (b) above is actually an equality.

A variant of part (b) will be needed in the infinite dimensional setting when the F-projection set is empty. If $\Pi_S^F(x) = \emptyset$, there are nonetheless points $s \in S$ that are suboptimal for (4) in the sense of (6) below. The content of the corollary is that from a fixed $x \notin S$, any point z on a line originating from x with the suboptimal velocity is suboptimal with the same error for the problem of minimizing x to the level set containing z. The corollary includes the optimal case $\varepsilon = 0$ as well.

Corollary 2.7 (Principle of Suboptimality). *Suppose $x \notin S$, $\varepsilon \geq 0$, and $s \in S$ satisfy*

$$\rho_F(s - x) \leq T(x) + \varepsilon. \tag{6}$$

Let $v := \frac{s-x}{\rho_F(s-x)} \in F$, and define $z_t := x + tv$ for $t \geq 0$. Now suppose $0 \leq r \leq T(x)$ and $\bar{t}$ satisfy $T(z_{\bar{t}}) = r$. Then

$$\bar{t} \leq \min_{z \in S(r)} \rho_F(z - x) + \varepsilon. \tag{7}$$

Proof. We have

$$\begin{aligned} r &= \min_{s' \in S} \rho_F(s' - z_{\bar{t}}) \leq \rho_F(s - z_{\bar{t}}) \\ &= \rho_F\left(s - x - \bar{t}\frac{s - x}{\rho_F(s - x)}\right) = \rho_F\left(\left[\rho_F(s - x) - \bar{t}\right]\frac{s - x}{\rho_F(s - x)}\right) \\ &= \rho_F(s - x) - \bar{t}. \end{aligned}$$

Thus by (6) we have

$$\bar{t} \leq \rho_F(s - x) - r \leq T(x) - r + \varepsilon,$$

and the final conclusion (7) follows from the previous Theorem, part (b), since it says $T(x) - r \leq \min_{z \in S(r)} \rho_F(z - x)$. □

3. General Formulas for $\partial_p T$ and $\partial_f T$

The first result in this section characterizes the proximal and the Fréchet subgradient of $T(\cdot)$ in general terms. The formulas have two features and can be naturally explained as follows: (1) one feature is to be expected from vector calculus in that the gradient is normal to the level set (although this is not true for general nonsmooth functions), and (2) the other says that the gradient is scaled in a manner to satisfy the Hamilton-Jacobi equation. Next, we prove an upper inclusion for both proximal and Fréchet subgradient at some point $x \notin S$, which does not involve the level sets of T. Reversing this inclusion, and hence explicitly computing the subgradients, will be one of the major efforts of Section 5.

Theorem 3.1. *Suppose S is closed and $F \subset H$ is closed, convex, bounded, and with $0 \in int\, F$. Suppose $x \notin S$ and $T(x) = r$. Then*

$$\partial_p T(x) = N^p_{S(r)}(x) \cap \{\zeta : \rho_{F^\circ}(-\zeta) = 1\},$$

$$\partial_f T(x) = N^f_{S(r)}(x) \cap \{\zeta : \rho_{F^\circ}(-\zeta) = 1\},$$

where $S(r)$ is as in (5).

Proof. We refer to [14] for the proof of the proximal case. Modifications of that argument will establish the formula for the Fréchet subgradient, which we now elucidate.

($\subseteq$) Let $\zeta \in \partial_f T(x)$. Then

$$\liminf_{y \to x} \frac{T(y) - T(x) - \langle \zeta, y - x \rangle}{\|y - x\|} \geq 0. \tag{8}$$

If we restrict the limit to $y \in S(r)$, the above formula becomes

$$\limsup_{y \to x,\, y \in S(r)} \frac{\langle \zeta, y - x \rangle}{\|y - x\|} \leq 0,$$

which says exactly that $\zeta \in N^f_{S(r)}(x)$.

We next show $\rho_{F^\circ}(-\zeta) \leq 1$. Let $v \in F$, and set, for $t \geq 0$, $y_t = x - tv$. Let $\varepsilon > 0$. If t is small enough, by (8) and Proposition 2.6 (a) we obtain that

$$\frac{1 + \langle \zeta, v \rangle}{\|v\|} \geq \frac{T(y_t) - T(x) - \langle \zeta, y_t - x \rangle}{\|y_t - x\|} \geq -\varepsilon.$$

By letting $\varepsilon \to 0$ we obtain

$$\rho_{F^\circ}(-\zeta) = \max_{v \in F} \langle -\zeta, v \rangle \leq 1. \tag{9}$$

Finally, we show there exists $\bar{v} \in F$ with $\langle \zeta, \bar{v} \rangle \leq -1$, which along with (9) implies $\rho_{F^\circ}(-\zeta) = 1$ as desired. For $t > 0$, let $s_t \in S$ be so that $\rho_F(s_t - x) \leq r + t^2$, and let $v_t = \frac{s_t - x}{\rho_F(s_t - x)} \in F$. Since F is weakly compact, there exists a sequence $\{t_i\}_i$ with $t_i \searrow 0$ and $v_i := v_{t_i}$ converging weakly to some $\bar{v} \in F$ as $i \to \infty$. Now consider $y_i := x + t_i v_i$, and write s_i for s_{t_i}. Observe that for all $t' > 0$

$$\frac{1}{t'}(s_i - y_i) = \frac{\rho_F(s_i - x) - t_i}{t'} \frac{s_i - x}{\rho_F(s_i - x)},$$

and so the minimum value of t' with $\frac{1}{t'}(s_i - y_i)$ belonging to F must necessarily satisfy $t' \le \rho_F(s_i - x) - t_i$. Therefore

$$\begin{aligned} T(y_i) = \min\{\rho_F(s - y_i) : s \in S\} &\le \min\{t' : \frac{1}{t'}(s_i - y_i) \in F\} \\ &\le \rho_F(s_i - x) - t_i < r - t_i + t_i^2 \end{aligned}$$

for all i. Let $\varepsilon > 0$. For large i, the previous estimate can be used in conjunction with (8) to obtain

$$\begin{aligned} r - t_i + t_i^2 > T(y_i) &\ge r + \langle \zeta, y_i - x\rangle - \varepsilon\|y_i - x\| \\ &= r + t_i\langle \zeta, v_i\rangle - \varepsilon t_i\|v_i\|. \end{aligned}$$

Now divide by $t_i > 0$ and let $i \to \infty$. Since $v_i \to \bar{v}$ weakly and ε is arbitrary, the conclusion is

$$\langle \zeta, \bar{v}\rangle \le -1. \tag{10}$$

Hence $\rho_{F^\circ}(-\zeta) = 1$ as asserted.

($\supseteq$) Now let ζ be such that $\rho_{F^\circ}(-\zeta) = 1$ and

$$\limsup_{z\to x,\, z\in S(r)} \frac{\langle \zeta, z - x\rangle}{\|z - x\|} \le 0. \tag{11}$$

Let $\varepsilon > 0$ be fixed. In order to prove (8) we must find $\delta > 0$ such that if $\|y - x\| \le \delta$ then

$$T(y) - r - \langle \zeta, y - x\rangle \ge -\varepsilon\|y - x\|. \tag{12}$$

There are three possibilities for a point y, which we shall consider separately: (i) $T(y) = r$, (ii) $T(y) > r$, and (iii) $T(y) < r$.

(i) The case $T(y) = r$ is trivial, since (12) follows automatically from (11).

(ii) Suppose $T(y) > r$ and $\|y - x\| \le 1$. There exists $s \in S$ such that

$$\rho_F(s - y) < T(y) + \|y - x\|^2. \tag{13}$$

Set $v := \frac{s-y}{\rho_F(s-y)}$, and choose $\bar{t}$ so that $z_t := y + tv$ satisfies $T(z_{\bar{t}}) = r$. Indeed, since $t_0 = 0$ satisfies $T(z_{t_0}) = T(y) > r$ and $t_1 = \rho_F(s - y)$ satisfies $T(z_{t_1}) = 0$, such a $\bar{t}$ exists by the intermediate value theorem. We claim that there exists a constant $k > 0$ independent of y such that

$$\|z_{\bar{t}} - x\| \le k\|y - x\| \tag{14}$$

Indeed, the Principle of Suboptimality (Corollary 2.7, with $\varepsilon = \|y - x\|^2$) implies that

$$\begin{aligned} \bar{t} &\le \min_{z\in S(r)} \rho_F(z - y) + \|y - x\|^2 \\ &\le \rho_F(x - y) + \|y - x\|^2 && \text{(since } x \in S(r)) \\ &\le (\|F^\circ\| + 1)\|y - x\|, && \text{(by (3) and } \|y - x\| \le 1) \end{aligned}$$

and thus

$$\begin{aligned}\|z_{\bar t}-x\| \le \|z_{\bar t}-y\|+\|y-x\| &= \bar t\|v\|+\|y-x\| \\ &\le \left\{(\|F^\circ\|+1)\|F\|+1\right\}\|y-x\| \\ &=: k\|y-x\|.\end{aligned}$$

Now we claim that

$$r+\bar t \le \rho_F(s-y). \tag{15}$$

Indeed,

$$s-z_{\bar t} = s-y-\bar t\frac{s-y}{\rho_F(s-y)} = \left(\rho_F(s-y)-\bar t\right)\frac{s-y}{\rho_F(s-y)},$$

and so

$$r=\min_{s'\in S}\rho_F(s'-z_{\bar t}) \le \rho_F(s-z_{\bar t}) \le \rho_F(s-y)-\bar t,$$

which implies (15).

If y is close enough to x, we can combine (13) and (15) and exploit (11) to obtain the following estimate

$$\begin{aligned}T(y)+\|y-x\|^2 &> r+\bar t \\ &\ge r+\bar t+\langle\zeta,z_{\bar t}-x\rangle-\varepsilon\|z_{\bar t}-x\| \\ &= r+\bar t+\bar t\langle\zeta,v\rangle+\langle\zeta,y-x\rangle-\varepsilon\|z_{\bar t}-x\| \\ &\ge r+\langle\zeta,y-x\rangle-\varepsilon\|z_{\bar t}-x\|,\end{aligned} \tag{16}$$

where we used the assumption $\rho_{F^\circ}(-\zeta)=1$ (which implies $\langle\zeta,v\rangle\ge -1$) to deduce the inequality in (16). Combining this estimate with (14) yields, for y close enough to x,

$$T(y)\ge r+\langle\zeta,y-x\rangle-2k\varepsilon\|y-x\|,$$

and finishes the proof of (ii).

(iii) Assume now that $T(y)<r$. Take $\overline{v}\in F$ such that $\langle\zeta,\overline{v}\rangle=-1$ and assume that $\varepsilon<1/(\|F^\circ\|\,\|F\|)$. Take $\delta>0$ so that $\|z-x\|<\delta$ and $z\in S(r)$ imply $\langle\zeta,z-x\rangle<\varepsilon\|\zeta\|\,\|z-x\|$. Set $z_t=y-t\overline{v}$. We claim that there exist $\bar t$ and k (the latter being independent of y) so that $z_{\bar t}\notin S(r)$ and $0\le\bar t\le k\|y-x\|$. Indeed, if y is close enough to x and t is small, then $z_t\in S(r)$ and

$$\langle\zeta,z_t-x\rangle<\varepsilon\|z_t-x\|\,\|\zeta\| \tag{17}$$

moreover,

$$\frac{\langle\zeta,z_t-x\rangle}{t}=1+\frac{\langle\zeta,y-x\rangle}{t}\to 1 \qquad \text{for } t\to+\infty.$$

On the other hand,

$$\lim_{t\to+\infty}\frac{\varepsilon\|z_t-x\|\,\|\zeta\|}{t}=\varepsilon\|\zeta\|\,\|\overline{v}\|\le\varepsilon\|F\|\,\|F^\circ\|<1.$$

Hence there exists $\hat t$ such that

$$\langle\zeta,z_{\hat t}-x\rangle=\varepsilon\|z_{\hat t}-x\|\,\|\zeta\|. \tag{18}$$

Thus

$$\begin{aligned}\hat{t} + \langle \zeta, y - x\rangle = \langle \zeta, z_{\hat{t}} - x\rangle &= \varepsilon \|z_{\hat{t}} - x\| \, \|\zeta\| \\ &\leq \varepsilon \|\zeta\| (\|y - x\| + \hat{t}\|\overline{v}\|) \\ &\leq \varepsilon \|F^\circ\| (\|y - x\| + \hat{t}\|F\|),\end{aligned}$$

from which one has

$$\hat{t}(1 - \varepsilon\|F^\circ\| \, \|F\|) \leq (\varepsilon + 1)\|F^\circ\| \, \|y - x\|$$

and

$$\hat{t} \leq k\|y - x\|, \qquad \text{with } k = \frac{(\varepsilon + 1)\|F^\circ\|}{1 - \varepsilon\|F^\circ\| \, \|F\|}.$$

If y is close enough to x, then $z_{\hat{t}} \notin S(r)$, because the condition (17) is violated. Now, since, for small t, $T(z_t) < r$, by the above argument we obtain that, for y close enough to x, there exists $0 < \bar{t} < \hat{t}$ such that

$$T(z_{\bar{t}}) = r, \qquad \bar{t} \leq k\|y - x\|, \tag{19}$$

and (17) holds.

We now return to proving (8). We have by the principle of optimality again that

$$T(z_{\bar{t}}) \leq T(y) + \bar{t},$$

from which it follows, recalling that $\langle \zeta, \overline{v}\rangle = -1$,

$$\begin{aligned}T(y) &\geq T(z_{\bar{t}}) - \bar{t} = r - \langle \zeta, -\bar{t}\overline{v}\rangle \\ &= r + \langle \zeta, y - x\rangle - \langle \zeta, z_{\bar{t}} - x\rangle.\end{aligned} \tag{20}$$

Now by (17)

$$\langle \zeta, z_{\bar{t}} - x\rangle < \varepsilon \|z_{\bar{t}} - x\| \, \|\zeta\|. \tag{21}$$

Also, from the $\bar{t}$-estimate in (19), we obtain

$$\|z_{\bar{t}} - x\| \leq \|y - x\| + \bar{t}\|\bar{v}\| \leq (1 + k\|F^\circ\|) \, \|y - x\|. \tag{22}$$

Substituting (21) and (22) into (20) yields (8). □

We state the following result as a corollary to the previous proof (see (10)), for it will be used subsequently.

Corollary 3.2. *Suppose $x \notin S$, $s \in \Pi_S^F(x)$, and $-\zeta \in \partial\rho_F(s - x)$. Then*

$$\left\langle \zeta, \frac{s - x}{\rho_F(s - x)} \right\rangle = 1.$$

The following result is a first step in order to give an alternative formula to those in Theorem 3.1, based on the F-projections rather than on the level sets of $T(\cdot)$.

Theorem 3.3. *Suppose $S \subseteq H$ is closed, $x \notin S$, and $\Pi_S^F(x) \neq \emptyset$. Then for all $s \in \Pi_S^F(x)$ the following inclusions hold:*

$$\partial_p T(x) \subseteq N_S^p(s) \cap (-\partial\rho_F(s - x)), \tag{23}$$

$$\partial_f T(x) \subseteq N_S^f(s) \cap (-\partial\rho_F(s - x)). \tag{24}$$

Proof. Fix $s \in \Pi_S^F(x)$. Of course both inclusions are trivial if $\partial_p T(x) = \partial_f T(x) = \emptyset$, so suppose $\zeta \in \partial_p T(x)$, and set $r := T(x)$. Recall Theorem 3.1, which says

$$\rho_{F^\circ}(-\zeta) = 1 \quad \text{and} \quad \zeta \in N^p_{S(r)}(x).$$

By Corollary 2.3, it suffices to show that

$$\zeta \in \left[-N_F\left(\frac{s-x}{\rho_F(s-x)}\right)\right] \cap N^p_S(s). \tag{25}$$

So, set $\bar{v} := \frac{s-x}{\rho_F(s-x)}$ and note by Corollary 3.2 that $\langle -\zeta, \bar{v}\rangle = 1$. Also, since $\rho_{F^\circ}(-\zeta) = 1$, we have $\langle -\zeta, v\rangle \le 1$ for all $v \in F$, and hence

$$\langle -\zeta, v - \bar{v}\rangle \le 0 \quad \forall v \in F,$$

which says that $\zeta \in -N_F(\bar{v})$.

We are left to showing that $\zeta \in N^p_S(s)$, and are assuming there exists $\sigma > 0$ so that

$$\langle \zeta, y - x\rangle \le \sigma \|y - x\|^2 \quad \forall y \in S(r). \tag{26}$$

Let $s' \in S$, and note that $y := s' + x - s$ belongs to $S(r)$ (since $T(y) \le \rho_F(s' - y) = \rho_F(s - x) = r$). Since $s' - y = s - x$ and $y - x = s' - s$, we have by (26) that

$$\begin{aligned} \langle \zeta, s' - s\rangle &= \langle \zeta, y - x\rangle \\ &\le \sigma \|y - x\|^2 \\ &= \sigma \|s' - s\|^2. \end{aligned}$$

Hence $\zeta \in N^p_S(s)$.

Let now $\zeta \in \partial_f T(x)$. The same argument as above shows that $\zeta \in -N_F(\bar{v})$. To prove that $\zeta \in N^f_S(s)$, take $s' \in S$ and set $y(s') = s' + x - s$. Observe as above that $y(s') \in S(r)$. Thus,

$$\limsup_{s' \to s} \frac{\langle \zeta, s' - s\rangle}{\|s' - s\|} = \limsup_{s' \to s} \frac{\langle \zeta, y(s') - x\rangle}{\|y(s') - x\|} \le 0,$$

which concludes the proof. □

4. The case where S is convex

As in many situations involving convexity, there is a global and complete characterization under additional convexity assumptions. We shall show in this section that the subgradient of $T(\cdot)$ at each point can be completely described when S is convex. It is convenient to have the following concept.

Definition 4.1. Suppose S is convex, $\bar{x} \notin S$, and $\bar{s} \in \Pi_S^F(\bar{x})$. The S/F separating normal cone SEP $(S/F, \bar{s}, \bar{x})$ for $(\bar{s}, \bar{x})$ is defined by

$$\text{SEP } (S/F, \bar{s}, \bar{x}) := N_S(\bar{s}) \cap \left\{-N_F\left(\frac{\bar{s} - \bar{x}}{\rho_F(\bar{s} - \bar{x})}\right)\right\}.$$

Theorem 4.2. *Suppose S is convex. Then*

(a) *$T(\cdot)$ is convex on H;*
(b) *for each $x \in H$, the F-projection set $\Pi_S^F(x)$ is not empty;*
(c) *for all $\bar{x} \notin S$, the separating cone $SEP\,(S/F, \bar{s}, \bar{x})$ is independent of the choice of $\bar{s} \in \Pi_S^F(\bar{x})$ and is nontrivial;*
(d) *the convex subgradient $\partial T(\bar{x})$ is given by*

$$\begin{aligned}\partial T(\bar{x}) &= \left\{ \frac{\zeta}{\rho_{F^\circ}(-\zeta)} : \zeta \in SEP\,(S/F, \bar{s}, \bar{x}) \text{ for some } \bar{s} \in \Pi_S^F(\bar{x}) \right\} \\ &= \left\{ \frac{\zeta}{\rho_{F^\circ}(-\zeta)} : \zeta \in SEP\,(S/F, \bar{s}, \bar{x}) \text{ for all } \bar{s} \in \Pi_S^F(\bar{x}) \right\}\end{aligned}$$

Proof. (a). Since $\rho_F(\cdot)$ is convex, the function $(x, s) \mapsto \rho_F(s - x) + I_S(s)$ is convex jointly in (x, s). The convexity of $T(\cdot)$ then follows directly from (4) and a general fact about minimizing over one of the variables a function that is jointly convex. We give a direct proof for the sake of completeness.

Let x_1, x_2 be elements in H, $0 \le \lambda \le 1$ and $\varepsilon > 0$. Let $t_i = T(x_i)$, $i = 1, 2$. Then

$$S \cap \big\{ x_1 + (t_1 + \varepsilon)F \big\} \neq \emptyset \quad \text{and} \quad S \cap \big\{ x_2 + (t_2 + \varepsilon)F \big\} \neq \emptyset.$$

Multiplying the first intersection by λ, the second by $(1 - \lambda)$, and adding, we get that the intersection of $S = \lambda S + (1 - \lambda)S$ with

$$\big\{ \lambda x_1 + (1 - \lambda)x_2 + \big(\lambda t_1 + (1 - \lambda)t_2 + \varepsilon\big)F \big\}$$

is not empty. We have used twice the fact that for any convex C and nonnegative numbers a and b, the distributive property $aC + bC = (a + b)C$ holds. Letting $\varepsilon \searrow 0$, we now conclude that

$$T(\lambda x_1 + (1 - \lambda)x_2) \le \lambda T(x_1) + (1 - \lambda)T(x_2),$$

or that $T(\cdot)$ is convex.

(b) For each $i = 1, 2 \ldots$, let $s_i \in S$ satisfy $\rho_F(s_i - x) < T(x) + \frac{1}{i}$. Since the sequence $\{s_i\}$ is bounded, there exists $\bar{s} \in H$ and a subsequence of $\{s_i\}$ that converges weakly to $\bar{s}$. But S and $x + T(x)F$ are both closed and convex and therefore weakly closed. Thus $\bar{s} \in S \cap \big\{ x + T(x)F \big\} = \Pi_S^F(x)$.

(c) This will follow from the more general statement:

Claim 4.3. *Suppose S_1 and S_2 are closed convex sets with S_2 bounded and $\operatorname{int} S_2 \neq \emptyset$. If $S_1 \cap \operatorname{int} S_2 = \emptyset$ and $S_1 \cap S_2 \neq \emptyset$, then for each $\bar{s} \in S_1 \cap S_2$, the cone $N_{S_1}(\bar{s}) \cap \{-N_{S_2}(\bar{s})\}$ is not the trivial cone $\{0\}$ and is independent of the particular choice of $\bar{s} \in S_1 \cap S_2$.*

Proof of Claim 4.3. The set $\tilde{S} := S_1 - S_2$ is convex, closed (since S_2 is bounded), and with nonempty interior. By assumption, $0 \in \operatorname{bdry} \tilde{S}$, and thus there exists $\zeta \in N_{\tilde{S}}(0)$ with $\zeta \neq 0$; see [15, Corollary 1.3, p. 5]. This means

$$\langle \zeta, s_1 - s_2 \rangle \le 0 \quad \forall s_1 \in S_1 \text{ and } s_2 \in S_2. \tag{27}$$

Now let $\bar{s} \in S_1 \cap S_2$. Setting $s_2 = \bar{s}$ in (27) illustrates that $\zeta \in N_{S_1}(\bar{s})$, while setting $s_1 = \bar{s}$ gives $-\zeta \in N_{S_2}(\bar{s})$. This proves the nontriviality statement. That any ζ does not depend on the choice of $\bar{s}$ follows directly as well since (27) does not depend on a particular choice of $\bar{s} \in S_1 \cap S_2$.

Part (c) of the Theorem follows from the claim by setting $S_1 = S$ and $S_2 = \bar{x} + T(\bar{x})F = \bar{x} + \rho_F(\bar{s} - \bar{x})F$ and noting that

$$N_{S_2}(\bar{s}) = N_F\left(\frac{\bar{s} - \bar{x}}{\rho_F(\bar{s} - \bar{x})}\right).$$

(d) Fix $\bar{x} \notin S$. By (c), the two sets on the right in (d) are equal. So for a fixed $\bar{s} \in \Pi_S^F(\bar{x})$ (which is nonempty by (b)), the conclusion of (d) is equivalent to the formula

$$\partial T(\bar{x}) = \left\{\frac{\zeta}{\rho_{F^\circ}(-\zeta)} : \zeta \in \text{SEP } (S/F, \bar{s}, \bar{x})\right\}. \tag{28}$$

To prove (28), set $r = T(\bar{x})$. Let $S(r)$ be as in (5), and observe that it is convex since $T(\cdot)$ is convex. Recall Theorem 3.1, which says

$$\partial T(\bar{x}) = N_{S(r)} \cap \{\zeta : \rho_{F^\circ}(-\zeta) = 1\}. \tag{29}$$

Comparing (28) and (29), it suffices to prove that

$$N_{S(r)}(\bar{x}) = \text{SEP } (S/F, \bar{s}, \bar{x}). \tag{30}$$

We break the equality in (30) into two inclusions.

($\supseteq$) Suppose $\zeta \in \text{SEP } (S/F, \bar{s}, \bar{x})$ and $x \in S(r)$. We must show that $\langle \zeta, x - \bar{x}\rangle \le 0$ for all $x \in S(r)$. The given vector ζ belongs to SEP $(S/F, \bar{s}, \bar{x})$, which implies two things:

$$\langle \zeta, s - \bar{s}\rangle \le 0 \qquad \forall s \in S, \qquad \text{and} \tag{31}$$

$$\left\langle \zeta, \frac{\bar{s} - \bar{x}}{r} - v\right\rangle \le 0 \qquad \forall v \in F \tag{32}$$

Let $x \in S(r)$ and $s \in \Pi_S^F(x)$, and note that

$$\rho_F(s - x) \le r = \rho_F(\bar{s} - \bar{x}). \tag{33}$$

We now calculate

$$\begin{aligned} \frac{1}{r}\langle \zeta, x - \bar{x}\rangle &= \frac{1}{r}\Big[\langle \zeta, x - s\rangle + \langle \zeta, s - \bar{s}\rangle + \langle \zeta, \bar{s} - \bar{x}\rangle\Big] \\ &\le \left\langle \zeta, \frac{\bar{s} - \bar{x}}{r} - \frac{s - x}{r}\right\rangle, \end{aligned}$$

where the inequality is valid by (31). Now $\frac{s-x}{\rho_F(s-x)} \in F$ and $\rho_F(s - x) \le r$ (by (33)), and since $0 \in F$ and F is convex, it follows that $v := \frac{s-x}{r} \in F$. Thus by (32), the last displayed line is nonpositive. Hence $\langle \zeta, x - \bar{x}\rangle \le 0$, as was to be shown.

($\subseteq$) Now let $\zeta \in N_{S(r)}(\bar{x})$, and so we have

$$\langle \zeta, x - \bar{x} \rangle \le 0 \qquad \forall x \in S(r) \tag{34}$$

and must show that (31) and (32) hold. Let $s \in S$, set $x := s + \bar{x} - \bar{s}$, and note that $x \in S(r)$ (since $\rho_F(s - x) = \rho_F(\bar{s} - \bar{x}) = r$). Hence by (34),

$$\begin{aligned} \langle \zeta, s - \bar{s} \rangle &= \langle \zeta, s - x \rangle + \langle \zeta, x - \bar{x} \rangle + \langle \zeta, \bar{x} - \bar{s} \rangle \\ &\le \langle \zeta, s - x \rangle + \langle \zeta, \bar{x} - \bar{s} \rangle \\ &= 0, \end{aligned}$$

and we have that (31) holds.

Now suppose $v \in F$, and let $x := \bar{s} - \rho_F(\bar{s} - \bar{x})v \in S(r)$. Then by (34) again,

$$\begin{aligned} \left\langle \zeta, \frac{\bar{s} - \bar{x}}{\rho_F(\bar{s} - \bar{x})} - v \right\rangle &= \left\langle \zeta, \frac{\bar{s} - \bar{x}}{\rho_F(\bar{s} - \bar{x})} - \frac{\bar{s} - x}{\rho_F(\bar{s} - \bar{x})} \right\rangle \\ &= \frac{1}{\rho_F(\bar{s} - \bar{x})} \langle \zeta, x - \bar{x} \rangle \\ &\le 0, \end{aligned}$$

and hence (32) holds. $\square$

As a restatement of the above result, we obtain that (if S is convex) the inclusions in Theorem 3.3 are actually equalities.

Corollary 4.4. *Suppose S is convex and $x \notin S$. Then*

$$\partial T(x) = N_S(s) \cap (-\partial \rho_F(s - x)) \quad \forall s \in \Pi_S^F(x).$$

Moreover, the set $N_S(s) \cap (-\partial \rho_F(s - x))$ is nonempty for each $s \in \Pi_S^F(x)$ and is independent of s.

If S is strictly convex, some regularity properties of the F-projection Π_S^F can be established (see Theorem 5.8 below). An example illustrating the above subgradient formula is contained in Section 6 below.

5. Results for nonconvex S

We now consider conditions on F and a possibly nonconvex target S for which $T(\cdot)$ has some regularity properties in a neighborhood of S. Both Fréchet and proximal subgradients are studied. The main tool for the present analysis is the projection map, the study of which is the subject of the first subsection.

In what follows we need a regularity concept for sets. The idea is that the nonconvexity of the target S has to be controlled and balanced by the strict convexity of the dynamics F.

Definition 5.1. A closed set $S \subset H$ is proximally smooth if there exists $\varphi \ge 0$ so that for all $s_1, s_2 \in S$ and $\zeta_i \in N_S^p(s_i)$ such that $\|\zeta_i\| \le 1$, $i = 1, 2$, one has

$$\langle \zeta_2 - \zeta_1, s_2 - s_1 \rangle \ge -\varphi \|s_2 - s_1\|^2. \tag{35}$$

To emphasize the dependence of φ, S is called φ-proximally smooth if (35) holds.

Obviously, a convex set is φ-proximally smooth with $\varphi = 0$, and it is easy to see that the boundary of the unit ball is 1-proximally smooth. It is also not difficult to show that if $S = \{x \in H : g(x) \leq 0\}$ for some $g \in \mathcal{C}^{1,+}(H)$ (i.e., g is differentiable and its differential Dg is Lipschitz continuous) and $Dg(x) \neq 0$ for all $x \in \text{bdry}\, S$, then S is φ-proximally smooth with φ the Lipschitz rank of Dg.

Definition 5.2. Let $\gamma > 0$ be given. We say that a closed convex set $F \subset H$ is γ-strictly convex if for all $v_1, v_2 \in F$ and $\zeta_i \in N_F(v_i)$, such that $\|\zeta_i\| \leq 1$, $i = 1, 2$, one has

$$\langle \zeta_2 - \zeta_1, v_2 - v_1 \rangle \geq \gamma \|v_2 - v_1\|^2. \tag{36}$$

The property of proximal smoothness has several consequences and admits various characterizations (see, e.g., [12, 16, 13, 6]). In particular, proximally smooth sets were studied in infinite dimensions in [12] in connection with the differentiability of the Euclidean distance in a neighborhood of S. We mention here some properties which will be needed in the proofs.

Proposition 5.3.

(*i*) *Let S be φ-proximally smooth, and let $s \in S$. Let $s_1, s_2 \in S$ be such that $\varphi\|s_i - s\| < 1$, $i = 1, 2$. Then for all $t \in [0, 1]$ one has*

$$d_S(ts_1 + (1-t)s_2) \leq \varphi t(1-t)\|s_1 - s_2\|^2; \tag{37}$$

moreover, for all $s \in S$, $N_S^p(s) = N_S^f(s) = N_S^l(s) = N_S^c(s)$. Furthermore, the normal cone has strongly$\times$weakly-closed graph, i.e., if $s_n \to s$, with $s_n \in S$, and $\zeta_n \to \zeta$ weakly, with $\zeta_n \in N_S^p(s_n)$, then $\zeta \in N_S^p(s)$.

(*ii*) *Let F be γ-strictly convex and let $v_1, v_2 \in$ bdry F. Then for all $t \in [0, 1]$ one has*

$$d_{bdry\, F}(tv_1 + (1-t)v_2) \geq \gamma t(1-t)\|v_1 - v_2\|^2. \tag{38}$$

Proof. (i) is Proposition 2.13 in [6] (with φ in place of $2p$, see also Remark 1.6 in [6]) and Propositions 5.1 and 6.1 in [13].

(ii). Fix $t \in (0, 1)$, set $v_t = tv_1 + (1-t)v_2$ and choose $\varepsilon > 0$. Since F has nonempty interior, by a density theorem (see [11, Theorem 1.3.1]) there exists $v_\varepsilon \in F$ such that $\|v_\varepsilon - v_t\| \leq \varepsilon$ and $\pi_{\text{bdry}\, F}(v_\varepsilon)$ is a singleton, say w_ε. We claim that $w_\varepsilon - v_\varepsilon \in N_F(w_\varepsilon)$. Indeed, if $w_\varepsilon = v_\varepsilon$ there is nothing to prove, otherwise take by contradiction $v \in F$ so that $\langle w_\varepsilon - v_\varepsilon, v - w_\varepsilon \rangle > 0$. Then the convex hull K of v and the ball $v_\varepsilon + \|w_\varepsilon - v_\varepsilon\| B$ contains w_ε in its interior. Since $K \subset F$ we reach a contradiction against $w_\varepsilon \in$ bdry F. Therefore, according to (36) one has both

$$\left\langle \frac{w_\varepsilon - v_\varepsilon}{\|w_\varepsilon - v_\varepsilon\|}, v_1 - w_\varepsilon \right\rangle \geq \gamma \|v_1 - w_\varepsilon\|^2, \qquad \left\langle \frac{w_\varepsilon - v_\varepsilon}{\|w_\varepsilon - v_\varepsilon\|}, v_2 - w_\varepsilon \right\rangle \geq \gamma \|v_2 - w_\varepsilon\|^2.$$

Multiplying by t the first inequality, by $1 - t$ the second one, and summing one obtains

$$\begin{aligned} \left\langle \frac{w_\varepsilon - v_\varepsilon}{\|w_\varepsilon - v_\varepsilon\|}, v_t - w_\varepsilon \right\rangle \ &\geq \ \gamma \left(t\|v_1 - w_\varepsilon\|^2 + (1-t)\|v_2 - w_\varepsilon\|^2 \right) = \\ &= \gamma \left(\|t(v_1 - w_\varepsilon) + (1-t)(v_2 - w_\varepsilon)\|^2 + t(1-t)\|v_1 - w_\varepsilon - (v_2 - w_\varepsilon)\|^2 \right). \end{aligned}$$

By the Cauchy-Schwartz inequality and rearranging the right-hand side, one has

$$\|v_t - w_\varepsilon\| \geq \gamma \left(\|v_t - w_\varepsilon\|^2 + t(1-t)\|v_1 - v_2\|^2 \right) \geq \gamma t(1-t)\|v_1 - v_2\|^2. \tag{39}$$

Recalling that $|d_{\text{bdry } F}(v_t) - d_{\text{bdry } F}(v_\varepsilon)| \leq \|v_t - v_\varepsilon\|$, we obtain $d_{\text{bdry } F}(v_t) \geq \|v_t - w_\varepsilon\| - 2\|v_t - v_\varepsilon\|$. By using (39) and passing to the limit for $\varepsilon \to 0$ we conclude the proof. □

The uniform strict convexity of F provides the following estimate for the gauge function.

Proposition 5.4. *Let F be γ-strictly convex. Then ρ_F satisfies the following estimate: for all x_1, x_2 one has*

$$\rho_F\left(\frac{x_1 + x_2}{2}\right) \leq \frac{\rho_F(x_1) + \rho_F(x_2)}{2} - \frac{\gamma}{2\|F\|} \frac{\rho_F(x_1)\rho_F(x_2)}{\rho_F(x_1) + \rho_F(x_2)} \left\| \frac{x_1}{\rho_F(x_1)} - \frac{x_2}{\rho_F(x_2)} \right\|^2. \tag{40}$$

Proof. Write $\rho_i = \rho_F(x_i)$ and $x_i = \rho_i v_i$ with $v_i \in \text{bdry } F$, $i = 1, 2$. Then

$$\frac{x_1 + x_2}{2} = \left(\frac{\rho_1}{\rho_1 + \rho_2} v_1 + \frac{\rho_2}{\rho_1 + \rho_2} v_2 \right) \frac{\rho_1 + \rho_2}{2}.$$

Set $v = (\rho_1 v_1 + \rho_2 v_2)/(\rho_1 + \rho_2)$, and observe that by, (38),

$$d_{\text{bdry } F}(v) \geq \gamma \frac{\rho_1 \rho_2}{(\rho_1 + \rho_2)^2} \|v_1 - v_2\|^2.$$

Assume $v \neq 0$. Then

$$\begin{aligned}
\rho_F\left(\frac{x_1 + x_2}{2}\right) &= \frac{\rho_1 + \rho_2}{2} \rho_F(v) \\
&= \frac{\rho_1 + \rho_2}{2} \left(1 - \rho_F\left(\frac{v}{\rho_F(v)} - v \right) \right) \\
&\leq \frac{\rho_1 + \rho_2}{2} - \frac{\rho_1 + \rho_2}{2\|F\|} \left\| \frac{v}{\rho_F(v)} - v \right\| \\
&\leq \frac{\rho_1 + \rho_2}{2} - \frac{\rho_1 + \rho_2}{2\|F\|} d_{\text{bdry } F}(v) \leq \frac{\rho_1 + \rho_2}{2} - \frac{\gamma}{2\|F\|} \frac{\rho_1 \rho_2}{\rho_1 + \rho_2} \|v_1 - v_2\|^2,
\end{aligned}$$

which is the desired estimate. If $v = 0$, then $x_1 + x_2 = 0$ as well, and it suffices to substitute in the above inequalities $v/\rho_F(v)$ with any point in bdry F. □

5.1. Regularity of the F-projection

In what follows, we adopt the convention that $\gamma/\varphi = \infty$ if $\varphi = 0$. The first result states that at an F-projection point a separation property, analogous to the nontriviality of SEP in the convex case, holds.

Proposition 5.5. *Let $S \subset H$ be closed and let $x \notin S$, $s \in \Pi_S^F(x)$. Then there exists $\zeta \in N_F\left((s - x)/\rho_F(s - x)\right) \cap \left(-N_S^l(s)\right)$, with $\zeta \neq 0$.*

Proof. Observe that s is a minimizer of the function $f(y) := \rho_F(y-x) + I_S(y)$. Thus, recalling [11, Proposition 10.1] and Corollary 2.3, one has

$$\begin{aligned} 0 &\in \partial_l\left(\rho_F(s-x) + I_S(s)\right) \subseteq \partial\rho_F(s-x) + N_S^l(s) \\ &= \{\zeta \in N_F\left((s-x)/\rho_F(s-x)\right) \,:\, \rho_{F^\circ}(\zeta) = 1\} + N_S^l(s) \end{aligned}$$

as asserted. □

The following result is the key point of our analysis. It states the existence and uniqueness of an optimal trajectory for a nonconvex target S, provided the starting point is close enough to S. Observe that this statement, and consequently all the subsequent analysis, does not contain the results from Section 4, since here we require the strict convexity of F.

Theorem 5.6. *Let S be φ-proximally smooth and F be γ-strictly uniformly convex, and let $x \in H$. Then:*

(*i*) *if $\varphi T(x) < \gamma$, then $\Pi_S^F(x)$ is at most a singleton;*
(*ii*) *if $\varphi T(x) < \frac{\gamma}{\|F^\circ\|\,\|F\|} \wedge \frac{1}{\|F\|}$, then $\Pi_S^F(x)$ is nonempty (and therefore it is a singleton).*

Proof. (i). Let $s_1, s_2 \in \Pi_S^F(x)$. By Proposition 5.5 there exists $\zeta \in N_F\left((s_1-x)\right.$ $\left./\rho_F(s_1-x)\right) \cap \left(-N_S^l(s_1)\right)$ with $\|\zeta\| = 1$. Set $\rho = \rho_F(s_1-x) = \rho_F(s_2-x)$. Then, by proximal smoothness,

$$\langle -\zeta, s_2 - s_1\rangle \le \varphi\|s_2 - s_1\|^2.$$

On the other hand, by strict convexity,

$$\langle -\zeta, \frac{s_2-s_1}{\rho}\rangle = \langle -\zeta, \frac{s_2-x}{\rho} - \frac{s_1-x}{\rho}\rangle \ge \frac{\gamma}{\rho^2}\|s_2-s_1\|^2.$$

Hence $\frac{\gamma}{\rho}\|s_2-s_2\|^2 \le \varphi\|s_2-s_1\|^2$. Thus if $\varphi\rho < \gamma$ one must have that $s_2 = s_1$.

(ii). Set $\rho = T(x)$, and let $\{s_n\} \subset S$ be a minimizing sequence for $\rho_F(\cdot - x)$, i.e., $\rho_F(s_n - x) := \rho_n \downarrow \rho$. We claim that $\{s_n\}$ is a Cauchy sequence. Indeed, by Proposition 5.4 one has

$$\rho_F\left(\frac{s_n+s_m}{2} - x\right) \le \frac{\rho_n+\rho_m}{2} - \frac{\gamma}{2\|F\|}\frac{\rho_n\rho_m}{\rho_n+\rho_m}\left\|\frac{s_n-x}{\rho_n} - \frac{s_m-x}{\rho_m}\right\|^2. \tag{41}$$

On the other hand, by taking $r = T(\frac{s_n+s_m}{2})$ in Proposition 2.6, part b), and recalling Proposition 2.5, (3), and (37), one obtains for m, n large enough

$$\rho_F\left(\frac{s_n+s_m}{2} - x\right) \ge \rho - T\left(\frac{s_n+s_m}{2}\right) \ge \rho - \frac{\varphi}{4}\|F^\circ\|\,\|s_n-s_m\|^2. \tag{42}$$

Fix now $\varepsilon > 0$. Observe that, since the sequence $\{s_n\}$ is bounded, it follows from (41) that for m, n large enough one has

$$\rho_F\left(\frac{s_n+s_m}{2} - x\right) \le \rho - \frac{\gamma}{4\rho\|F\|}\|s_n-s_m\|^2 + \varepsilon. \tag{43}$$

From (42) and (43) it then follows, for m, n large enough,

$$\left(\frac{\gamma}{4\rho\|F\|} - \frac{\varphi}{4}\|F^\circ\| \right) \|s_n - s_m\|^2 \leq \varepsilon,$$

which is the desired estimate. □

Theorem 5.7. *Let S be φ-proximally smooth and let F be γ-strictly convex. Let $\bar{r}$ be such that $0 < \varphi\bar{r} < \frac{\gamma}{\|F^\circ\|\,\|F\|} \wedge \frac{1}{\|F\|}$.*

(i) *Then the single valued map Π_S^F is locally Hölder continuous on $S(\bar{r})$ with exponent 1/2.*

(ii) *Let, furthermore, $F = \{v \in H : g(v) \leq 0\}$ for some $g \in \mathcal{C}^{1,+}(H)$ such that $Dg(v) \neq 0$ for all $v \in bdry\, F$. Alternatively, assume that $S = \{x \in H : f(x) \leq 0\}$ for some $f \in \mathcal{C}^{1,+}$ with $Df(s) \neq 0$ for all $s \in bdry\, S$. Then Π_S^F is locally Lipschitz continuous on $S(\bar{r})$.*

Proof. Theorem 5.6 (ii) yields that the F-projection is a singleton for all $x \in S(\bar{r})$. Set $s_i = \Pi_S^F(x_i)$ and $\rho_i = \rho_F(s_i - x_i)$, $i = 1, 2$.

Ad (i). The proof is divided into two steps.

Step 1: $\rho_1 = \rho_2 = \rho$.

Take $\zeta_i \in N_F\left((s_i - x_i)/\rho\right) \cap (-N_S^p(s_i))$ with $\|\zeta_i\| = 1$. Then, by proximal smoothness

$$\langle \zeta_2 - \zeta_1, s_2 - s_1 \rangle \leq \varphi \|s_2 - s_1\|^2, \tag{44}$$

while by strict convexity and Cauchy–Schwartz inequality

$$\begin{aligned} \langle \zeta_2 - \zeta_1, \frac{s_2 - x_2}{\rho} - \frac{s_1 - x_1}{\rho} \rangle &\geq \frac{\gamma}{\rho^2}\|s_2 - x_2 - (s_1 - x_1)\|^2 \geq \\ &\geq \frac{\gamma}{\rho^2}\left(\|s_2 - s_1\|^2 + \|x_2 - x_1\|^2 - 2\|s_2 - s_1\|\,\|x_2 - x_1\|\right). \end{aligned} \tag{45}$$

Putting together (44) and (45) and using again Cauchy–Schwartz inequality, one obtains

$$\frac{\gamma}{\rho}\left(\|s_2 - s_1\|^2 + \|x_2 - x_1\|^2 - 2\|s_2 - s_1\|\,\|x_2 - x_1\|\right) \leq \varphi\|s_2 - s_1\|^2 + \|\zeta_2 - \zeta_1\|\,\|x_2 - x_1\|, \tag{46}$$

which implies

$$(\gamma - \rho\varphi)\|s_2 - s_1\|^2 - 2\gamma\|x_2 - x_1\|\,\|s_2 - s_1\| - 2\rho\|x_2 - x_1\| \leq 0. \tag{47}$$

Therefore

$$\|s_2 - s_1\| \leq \frac{\gamma\|x_2 - x_1\| + \sqrt{\gamma^2\|x_2 - x_1\| + 2\rho(\gamma - \rho\varphi)}\sqrt{\|x_2 - x_1\|}}{\gamma - \rho\varphi},$$

which implies the Hölder continuity. Observe that the Hölder ratio, which we denote by H_ρ, depends on the level ρ.

Step 2: ρ_1 and ρ_2 arbitrary.

We take $\rho_2 > \rho_1$, the other case being handled symmetrically. Recall first that $|\rho_2 - \rho_1| \leq \|F^\circ\| \, \|x_2 - x_1\|$. Write $s_2 = x_2 + \rho_2 v$, with $\rho_F(v) = 1$, and let $x_2' = x_2 + (\rho_2 - \rho_1)v$. Observe that by the optimality principle (Proposition 2.6, part (a)) we have that $T(x_2') = \rho_1$ and $\Pi_S^F(x_2') = s_2$. Then

$$\begin{aligned} \|s_2 - s_1\| &\leq H_{\rho_1}\sqrt{\|x_2' - x_1\|} \leq H_{\rho_1}\sqrt{\|x_2' - x_2\| + \|x_2 - x_1\|} \\ &\leq H_{\rho_1}\sqrt{1 + \|F\| \, \|F^\circ\|}\sqrt{\|x_2 - x_1\|}, \end{aligned}$$

which concludes the proof of part (i).

Ad (ii). Assume again $\rho_1 = \rho_2 = \rho$, and denote by λ the Lipschitz ratio of Dg, and set $M = \sup\{\|Dg(v)\| \, : \, v \in F\} < \infty$. If F is $\mathcal{C}^{1,+}$, the normal vectors ζ_1, ζ_2 can be chosen to be, respectively, $Dg((s_1 - x_1)/\rho)/M$, $Dg((s_2 - x_2)/\rho)/M$. Then (36) yields

$$\frac{\lambda}{M\rho}\|s_2 - x_2 - (s_1 - x_1)\|^2 \geq \tag{48}$$

$$\geq \frac{1}{M}\left\langle Dg\left(\frac{s_2 - x_2}{\rho}\right) - Dg\left(\frac{s_1 - x_1}{\rho}\right), s_2 - x_2 - (s_1 - x_1)\right\rangle \tag{49}$$

$$\geq \frac{\gamma}{\rho}\|s_2 - x_2 - (s_1 - x_1)\|^2,$$

which, in particular, implies that $\gamma \leq \lambda/M$. By the inequality between (48) and (49), the right-hand side of (46) can be refined to be

$$\varphi\|s_2 - s_1\|^2 + \frac{\lambda}{M\rho}\left(\|s_2 - s_1\| + \|x_2 - x_1\|\right)\|x_2 - x_1\|.$$

Then (47) becomes

$$(\gamma - \rho\varphi)\|s_2 - s_1\|^2 - (2\gamma + \frac{\lambda}{M})\|x_2 - x_1\| \, \|s_2 - s_1\| + (\gamma - \frac{\lambda}{M})\|x_2 - x_1\|^2 \leq 0.$$

This yields the local Lipschitz estimate for $\rho_1 = \rho_2 = \rho$.

If S is of class $\mathcal{C}^{1,+}$, assume that $\|x_1 - \bar{s}\| \leq \eta$, $\|x_2 - \bar{s}\| \leq \eta$ for some $\bar{s} \in S$ and $\eta > 0$, set $K = \sup\{\|Df(s)\| \, : \, s \in \text{bdry}\, S, \, \|s - \bar{s}\| \leq 2\eta\}$ and let μ be the Lipschitz ratio of Df. Then we can set $\zeta_i = Df(s_i)/K$, $i = 1, 2$. The righthand side of (46) can be refined to be

$$\varphi\|s_2 - s_1\|^2 + \frac{\mu}{K}\|s_2 - s_1\| \, \|x_2 - x_1\|,$$

which implies

$$(\gamma - \rho\varphi)\|s_2 - s_1\|^2 - (2\gamma + \frac{\mu}{K})\|x_2 - x_1\| \, \|s_2 - s_1\| \leq 0.$$

The case of $\rho_1 = \rho_2$ arbitrary is handled as in Step 2 above. $\square$

The last result of the section concerns convex targets.

Theorem 5.8. *Let S be γ-strictly convex. Then* (*i*) *the projection map* $\Pi_S^F : H \to S$ *is a singleton which is Hölder continuous with exponent* $1/2$;

(*ii*) *if moreover* $S = \{x \in H : f(x) \le 0\}$ *for some* $f \in \mathcal{C}^{1,+}(H)$ *with* $Df(x) \ne 0$ *for all* $x \in bdry\, S$, *and* S *is bounded, then* Π_S^F *is Lipschitz continuous.*

Proof. Nonemptiness and uniqueness of the F-projection follow from standard arguments.

Ad (i). Take $x_1, x_2 \in H$, assuming first that $T(x_1) = T(x_2) = \rho$. Write $s_i = \Pi_S^F(x_i)$, $i = 1, 2$, and take by Proposition 5.5 $\zeta_i \in N_F((s_i - x_i)/\rho) \cap (-N_S(s_i))$, with $\|\zeta_i\| \le 1$. By the strict convexity of S one has

$$\langle -\zeta_2 + \zeta_1, s_2 - s_1 \rangle \ge \gamma \|s_2 - s_1\|^2.$$

On the other hand, by the convexity of F one has

$$\langle -\zeta_2 + \zeta_1, s_2 - s_1 \rangle \le \langle -\zeta_2 + \zeta_1, x_2 - x_1 \rangle.$$

Thus

$$\gamma \|s_2 - s_1\|^2 \le \|\zeta_2 - \zeta_1\| \, \|x_2 - x_1\|, \tag{50}$$

which gives the Hölder estimate. Observe that the Hölder ratio is independent of ρ. The general case is treated as in Step 2 of the proof of Theorem 5.7.

Ad (ii). If S is $\mathcal{C}^{1,+}$, by the same argument as in the proof of part (ii) of Theorem 5.7, the right-hand side of (50) can be refined to be

$$\frac{\mu}{K} \|s_2 - s_1\| \, \|x_2 - x_1\|,$$

where μ is the Lipschitz ratio of Df and $K = \sup\{\|Df(s)\| : s \in S\}$. The Lipschitz estimate then follows immediately. □

5.2. Fréchet and proximal subgradient

Our goal is to identify hypotheses so that the opposite inclusions in Theorem 3.3 hold, and thereby obtain regularity results for $T(\cdot)$. This is well-understood for the case of $F = \overline{I\!B}$ (see, e.g., [12, 6, 16]), and we state as corollaries a few situations where those results can be generalized. By Theorem 3.3, it is clear that if one seeks $\partial_p T(x) \ne \emptyset$ for x in a neighborhood of S, then S must have plentiful proximal normal cone properties, and proximal smoothness is such a condition.

We begin with the Fréchet subdifferential.

Proposition 5.9. *Let* S *be* φ*-proximally smooth and* F *be* γ*-strictly convex. Let* $x \in H \backslash S$ *and set* $r := T(x)$. *Assume that* $\varphi r < \frac{\gamma}{\|F^\circ\| \|F\|} \wedge \frac{1}{\|F\|}$. *Then*

$$\emptyset \ne N_S^p(\Pi_S^F(x)) \cap \left(-\partial \rho_F(\Pi_S^F(x) - x)\right) \subseteq N_{S(r)}^f(x). \tag{51}$$

Proof. Recalling Propositions 2.2, 5.5 and part (i) in 5.3, we obtain that the lefthand side of (51) is nonempty. Set now $s = \Pi_S^F(x)$. Let $\zeta \in -\partial \rho_F(s - x) \cap N_S^p(s)$, and let $x_n \to x$, $x_n \in S(r)$. We want to show that

$$\limsup_{n \to \infty} \left\langle \zeta, \frac{x_n - x}{\|x_n - x\|} \right\rangle \le 0. \tag{52}$$

To this aim, set $s_n = \Pi_S^F(x_n)$. Then, by the assumptions on S and F one has

$$\begin{aligned}\langle \zeta, x_n - x\rangle &= \langle \zeta, x_n - s_n\rangle + \langle \zeta, s_n - s\rangle + \langle \zeta, s - x\rangle \\ &= r\left\langle \zeta, \frac{s-x}{r} - \frac{s_n - x_n}{r}\right\rangle + \langle \zeta, s_n - s\rangle \\ &\le -\frac{\gamma\|\zeta\|}{r}\|s_n - x_n - (s-x)\|^2 + \varphi\|\zeta\|\,\|s_n - s\|^2 \\ &= \left(\varphi - \frac{\gamma}{r}\right)\|\zeta\|\,\|s_n - s\|^2 + \frac{2\gamma\|\zeta\|}{r}\langle s_n - s, x_n - x\rangle - \frac{\gamma}{r}\|\zeta\|\,\|x_n - x\|^2.\end{aligned}$$

Since $\|s_n - s\| = O(\sqrt{\|x_n - x\|})$ by Theorem 5.7 part (i), we obtain (52). □

As a consequence, we obtain the following result.

Theorem 5.10. *Let S be φ-proximally smooth and F be γ-strictly convex. Let $\bar{r}$ be such that $0 < \varphi\bar{r} < \frac{\gamma}{\|F^\circ\|\,\|F\|} \wedge \frac{1}{\|F\|}$. Let $x \in S(\bar{r}) \setminus S$, and set $s = \Pi_S^F(x)$. Then*

$$\partial_f T(x) = N_S^p(s) \cap (-\partial\rho_F(s - x)) \neq \emptyset. \tag{53}$$

Moreover, T is Clarke regular on $S(\bar{r}) \setminus S$ (i.e., the Clarke generalized gradient and the Fréchet subgradient coincide).

Proof. Since by Theorem 5.6 (ii) the F-projection of x is a singleton, formula (53) comes by putting together (51) and (24). To show the Clarke regularity, consider the following chain of inclusions:

$$\begin{aligned}\partial_f T(x) &= N^f_{S(r)}(x) \cap \{\zeta : \rho_{F^\circ}(-\zeta) = 1\} &(54)\\ &\subseteq N^c_{S(r)}(x) \cap \{\zeta : \rho_{F^\circ}(-\zeta) = 1\} &(55)\\ &= \overline{\mathrm{co}}\,\{\zeta : \zeta = w - \lim \zeta_i,\ \zeta_i \in \partial_f T(x_i),\ x_i \to x,\ x_i \in S(r)\} &(56)\\ &\subseteq \overline{\mathrm{co}}\,\{\zeta : \zeta = w - \lim \zeta_i,\ \zeta_i \in \partial_f T(x_i),\ x_i \to x\} &(57)\\ &= \overline{\mathrm{co}}\left\{\zeta : \zeta = w - \lim \zeta_i,\ \zeta_i \in N_S^p(\Pi_S^F(x_i)) \cap \left(-\partial\rho_F(\Pi_S^F(x_i) - x_i)\right),\ x_i \to x\right\} &(58)\\ &\subseteq \overline{\mathrm{co}}\,\{\zeta : \zeta \in N_S^p(s) \cap (-\partial\rho_F(s - x))\} &(59)\\ &= -\partial\rho_F(s - x) \cap N_S^p(s) = \partial_f T(x). &(60)\end{aligned}$$

Indeed, (54) is Theorem 3.1, the equality between (55) and (56) follows from [2, Theorem 4.4.4] and Theorem 3.1, while the inclusion between (58) and (59) follows from the normal regularity of S (see (i) in Proposition 5.3) and F and the continuity of the F-projection; the equality between (59) and (60) follows again from the normal regularity of S and F. Observe that the set in (57) is exactly the Clarke generalized gradient of $T(x)$, which therefore coincides with the Fréchet subgradient. The proof is concluded. □

Corollary 5.11. *Under the assumptions of Theorem* 5.10*, suppose that the righthand side of* (53) *is a singleton for some $x \in S(\bar{r})$. Then T is strictly differentiable at x (see* [10, p. 30]*).*

Proof. By Theorem 5.10, the righthand side of (53) is also the Clarke generalized gradient of T at x. By Proposition 2.2.4 in [10], T is strictly differentiable at x. $\square$

In addition to proximal smoothness, the following theorem hypothesizes a sort of one-sided Lipschitz condition of the F-projection map (see (61) below), and, just as in the convex case, concludes that (25) holds as an equality. It is not clear if (61) always holds for S convex. However, in the previous section we showed that in some cases the projection map is singleton-valued and Lipschitz, which immediately implies (61).

Theorem 5.12. *Let S be φ-proximally smooth, and suppose $x \notin S$ is such that $\Pi_S^F(x) \neq \emptyset$, and there exist constants $\eta = \eta(x) > 0$ and $k = k(x) > 0$ so that*

$$\Pi_S^F(y) \subseteq \Pi_S^F(x) + k\|y - x\|\overline{I\!B} \quad \forall y \in x + \eta\overline{I\!B}, \tag{61}$$

and that the set $N_S^p(s) \cap (-\partial\rho_F(s - x))$ is independent of $s \in \Pi_S^F(x)$. Then one has

$$\partial_p T(x) = N_S^p(s) \cap (-\partial\rho_F(s - x)) \neq \emptyset \tag{62}$$

for each $s \in \Pi_S^F(x)$.

Proof. The inclusion "$\subseteq$" is the result of Theorem 3.3. In view of Theorem 3.1, the opposite inclusion "$\supseteq$" follows if it can be shown that

$$\emptyset \neq N_S^p(s) \cap (-\partial\rho_F(s - x)) \subseteq N_{S(r)}^p(x) \tag{63}$$

for all $s \in \Pi_S^F(x)$, where $r := T(x)$. The nonemptiness of the lefthand side of (63) can be proved as in Proposition 5.9, so suppose ζ belongs to the left side of (63) for some (and therefore all) $s \in \Pi_S^F(x)$. Now let $y \in \left(x + \eta\overline{I\!B}\right) \cap S(r)$, and select any $s' \in \Pi_S^F(y)$. By (61), there exists $s \in \Pi_S^F(x)$ so that

$$\|s' - s\| \leq k\|y - x\|. \tag{64}$$

Now we write

$$\langle \zeta, y - x\rangle = r\left\langle -\zeta, \frac{s' - y}{r} - \frac{s - x}{r}\right\rangle + \langle \zeta, s' - s\rangle, \tag{65}$$

and note that $\rho_F(s' - y) = T(y) \leq r$ implies $\rho_F\left(\frac{s'-y}{r}\right) \leq 1$, or that $\frac{s'-y}{r}$ belongs to F. Since, by Proposition 2.3, $-\zeta \in N_F\left(\frac{s-x}{r}\right)$, the first term on the righthand side of (65) is thus nonpositive. The second term is bounded by $\varphi\|\zeta\|\,\|s' - s\|^2$ by (35), and so by (64) and (65), we have

$$\langle \zeta, y - x\rangle \leq \varphi k^2\|\zeta\|\,\|y - x\|^2.$$

This says $\zeta \in N_{S(r)}^p(x)$ and finishes the proof of (63), and consequently of the theorem. $\square$

5.3. Semiconvexity and differentiability

In this section we consider some simple consequences of our previous analysis. We prove first a result concerning a sufficient condition for semiconvexity (which is called lower $\mathcal{C}^2$ in [17], [12]) of $T(\cdot)$ near S. We recall that a function f from a convex set $U \subset H$ into $\mathbb{R}$ is called *semiconvex* if there exists a constant $C > 0$ such that for all $x_1, x_2 \in U$ and for all $\lambda \in [0, 1]$ one has

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) + \lambda(1 - \lambda)C\|x_1 - x_2\|^2.$$

Proposition 5.13. *Suppose S is closed, and let $U \subset H \setminus S$ be an open bounded convex set. Assume that the F-projection Π_S^F of each point x in U admits a single valued Lipschitz selection Π (as it is the case, for example, if the assumptions of part (ii) in Theorem 5.7 are satisfied), and assume, furthermore, that there exists $\varphi \geq 0$ such that the estimate (35) holds for all $s_i = \Pi(x_i)$, $x_i \in U$, and $\zeta_i \in N_S^p(s_i)$, $i = 1, 2$ (as it is the case, for example, if S is φ-proximally smooth). Then $\partial_p T(x) \neq \emptyset$ for all $x \in U$ and $T(\cdot)$ is semiconvex on U.*

Proof. Let $x_1, x_2 \in U$. Set $x = \lambda x_1 + (1-\lambda)x_2$ and choose $y_1 \in \Pi_S^F(x_1)$ and $y_2 \in \Pi_S^F(x_2)$ such that $\|y_1 - y_2\| \leq L\|x_1 - x_2\|$. Set $y = \lambda y_1 + (1-\lambda)y_2$. Then, by Proposition 2.6 (b), the convexity of ρ_F, (3), and (37), one has

$$\begin{aligned} T(x) &\leq \rho_F(y-x) + T(y) \\ &\leq \lambda\rho_F(y_1 - x_1) + (1-\lambda)\rho_F(y_2 - x_2) + d_S(y)\|F^\circ\| \\ &\leq \lambda T(x_1) + (1-\lambda)T(x_2) + \kappa\lambda(1-\lambda)L\|F^\circ\|\,\|x_1 - x_2\|^2, \end{aligned}$$

for a suitable constant $\kappa \geq 0$. The proof is concluded. □

The second result concerns sufficient conditions for the smoothness of T.

Theorem 5.14. *Let S and F satisfy the assumptions of Theorem 5.10, and assume furthermore that either F is the closure of an open convex set of class $\mathcal{C}^1$, or $S = \{s \in H : f(s) \leq 0\}$ for some $f \in \mathcal{C}^1(H)$ such that $Df(s) \neq 0$ for all $s \in$ bdry S. Let $\bar{r}$ be such that $0 < \varphi\bar{r} < \frac{\gamma}{\|F^\circ\|\,\|F\|} \wedge \frac{1}{\|F\|}$. Then T is of class $\mathcal{C}^1$ on $S(\bar{r}) \setminus S$, and*

$$DT(x) = N_S^p(\Pi_S^F(x)) \cap \left(-\partial\rho_F(\Pi_S^F(x) - x)\right). \tag{66}$$

Proof. By Theorem 5.10, the Fréchet subdifferential of T is computed as

$$\partial_f T(x) = N_S^p(\Pi_S^F(x)) \cap \left(-\partial\rho_F(\Pi_S^F(x) - x)\right)$$

for all $x \in S(\bar{r}) \setminus S$. By Proposition 5.5, the righthand side of (66) is nonempty. Since either ρ_F or S are of class $\mathcal{C}^1$, it is also a singleton, continuously depending on x, since Π_S^F is so. □

Remark. Observe that if in Theorem 5.14 F is the closure of an open $\mathcal{C}^{1,+}$ convex set, then (66) becomes

$$DT(x) = -D\rho_F(\Pi_S^F(x) - x)$$

This generalizes the well known formula for the gradient of the Euclidean distance from a proximally smooth set (see, e.g., [12]), and implies by Theorem 5.7 that $T(\cdot)$ is of class $\mathcal{C}^{1,+}$.

6. Examples

1) To illustrate Theorem 4.2, let F be the L^1-ball $\{(x,y) : |x| + |y| \leq 1\}$ in $\mathbb{R}^2$, and let $S = \{(x,y) : x, y \leq 0\}$. Our formula for $\partial T(\cdot)$ detects exactly the nondifferentiability points of T, which are the points of the positive x and y-axis. For example, the subgradient of T at $(0,1)$ is the set $\{(x,1) : 0 \leq x \leq 1\}$.

2) Let S be the convex hull of $\{(x,y,z) : x = 0,\ -1 \le y \le 1,\ z \ge 0\}$ and the curve $\{(x,y,z) : y = 0,\ z = x^2\}$, i.e., S is the "wall" defined by the first set, all the lines connecting the wall to the curve, and all the space above these lines. Thus each point on the curve is an extreme point. Now let F be the unit cube (i.e., the L^∞ unit ball), and consider points on the curve $\{(x,y,z) : x = r+1,\ y = 0,\ z = r^2 - 1,\ r \ge 0\}$. The latter curve is just the curve used in forming S except that it is shifted over and down by one unit. All the points on the curve are in the 1-level set of $T(\cdot)$. When $r = 0$, the F-projection is $\{(x,y,z) : x = 0,\ -1 \le y \le 1,\ z = 0\}$, and when $r > 0$, the projection is $\{(r+1, 0, r^2)\}$.

The above example shows that the projection map itself may not be continuous, even in the convex case. However, observe that in this case the projection map admits a Lipschitz continuous selection.

3) The property (61) may hold without the existence of a Lipschitz selection from Π_S^F. It suffices to take as F the product of the interval $[-1,1]$ with the L^1-ball in the (y,z)-space, and as S the convex hull of the sets $\{(x,0,0) : -1 \le x \le 1\}$, $\{(1,y,y^2) : y \le 0\}$ and $\{(-1,y,y^2) : y \ge 0\}$.

4) To illustrate formula (62) and the subsequent analysis, let F be $\{-\frac{1}{9} \le x \le \frac{1}{9},\ -10x - 1 \le y \le 10x - 1\}$ and $S = \{x \le 0,\ y \le x^2\}$. Observe that $-F \ne F$, so that ρ_F is not a norm, and that S is proximally smooth. Consider the 1-level set $S(1)$ of T, which is a parabola for $x \le 0$, while for $x \ge 0$ it is the segment joining $(0,1)$ with $(\frac{1}{9}, -\frac{1}{9})$ and then a vertical half line. It is easy to show that if $x \ge -1/2$, then $\Pi_S^F(x)$ is a singleton, which is Lipschitz w.r.t. x. By applying formula (62) it is easy to see that the only points of nondifferentiability in $S(1)$ for T are $(0,1)$ and $(\frac{1}{9}, -\frac{1}{9})$. If instead F is an ellipse, then T is $\mathcal{C}^{1,+}$ in a neighborhood of S, which can be explicitly determined.

5) The minimal time function needs not be semiconvex, even if the F-projection is unique. It suffices to take as F the L^1-ball in $\mathbb{R}^2$, and as S the set $\{y \ge f(x)\}$, with $f(x) = \min\{x/2, 0\}$.

6) Let $H = \mathbb{R}^2$, $S = \overline{B}_\infty$, $F = \overline{B}_1$. Let $x_0 = (1,2)$. Then Π_S^F is a singleton in a neighborhood of x_0, but the convex subgradient of T at x_0 is multivalued.

7) The assumptions of proximal smoothness in Theorem 5.7 and, very likely, in Theorem 5.8 are not sharp. Indeed, set $F := \{(x,y) : |y| \le 1 - |x|^{3/2}\}$ and $S := \{(x,y) : -2 \le y \le |x|^{3/2} - 1,\ |x| \le 1\}$, and observe that neither F nor S are φ-proximally smooth for any φ. However, for any x small enough, $\Pi_S^F(x, -1/2)$ is a singleton, Lipschitz continuous with respect to x. In fact, the unique point $(\xi(x), \eta(x))$ in $\Pi_S^F(x,-1/2)$ can be computed by imposing that the tangent line to S at $(\xi(x), \eta(x))$ is orthogonal to a normal to $(x,-1/2) + T(x,-1/2)F$ at $(\xi(x), \eta(x))$. This condition reads, for $x > 0$ small enough,

$$\frac{\sqrt{\xi(x) - x}}{\sqrt{T(x,-1/2)}} = \sqrt{\xi(x)},$$

which yields

$$\xi(x) = \frac{x}{1 - T(x,-1/2)}.$$

Since $T(x,-1/2) \to 1/2$ as $x \to 0^+$, we have that $\xi(\cdot)$ is Lipschitz continuous in a neighborhood of 0. It is easy to see that also $\eta(\cdot)$ is Lipschitz.

8) Here a non convex example is presented where property (61) holds without the uniqueness of the projection and the proximal smoothness of the target S. Let $F = \overline{I\!B}$ and $S = \{(x, y) : -1 \leq y \leq |x|^{3/2}, |x| \leq 1\}$. For $0 < y \leq 1$ the F-projection $\Pi_S^F(0, y)$ is never a singleton (actually it is $(\xi(y), \pm\eta(y))$, with $\eta(y) > 0$). Observe that $\|\Pi_S^F(0, y)\| \to 0$ as $y \to 0^+$. We compute $(\xi(y), \eta(y))$ in terms of y through $T(0, y)$, and show that they are Lipschitz continuous functions of y, at least for y small. Indeed, imposing the normality condition between the tangent to S and the normal to $(0, y) + T(0, y)F$ at $(\xi(y), \eta(y))$ gives

$$\frac{3}{2}\sqrt{\xi(y)} = \frac{\xi(y)}{\sqrt{T^2(0, y) - \xi^2(y)}},$$

which implies

$$\xi(y) = \frac{-2 + \sqrt{4 + 81T^2(0, y)}}{9}.$$

Being T Lipschitz, ξ and η are also Lipschitz.

Acknowledgements. The proof of Theorem 5.7 profited of an useful discussion with R. Vinter.

References

[1] P. Albano, P. Cannarsa, C. Sinestrari: Regularity results for the minimum time function of a class of semilinear evolution equations of parabolic type, SIAM J. Control Optimization 38 (2000) 916–946.

[2] J. P. Aubin, H. Frankowska: Set-Valued Analysis, Birkhäuser, Boston (1990).

[3] M. Bardi: A boundary value problem for the minimal time problem, SIAM J. Control Optimization 27 (1989) 776–785.

[4] M. Bounkhel, L. Thibault: On various notions of regularity of sets in nonsmooth analysis, Nonlinear Anal., TMA 48 (2002) 223–246.

[5] A. Bressan: On two conjectures by Hájek, Funkc. Ekv. 23 (1980) 221–227.

[6] A. Canino: On p-convex sets and geodesics, J. Diff. Equat. 75 (1988) 118–157.

[7] P. Cannarsa, R. Peirone: Unbounded components of the singular set of the distance function in $I\!R^n$, Trans. Amer. Math. Soc. 353 (2001) 4567–4581.

[8] P. Cannarsa, C. Sinestrari: Convexity properties of the minimum time function, Calc. Var. 3 (1995) 273–298.

[9] P. Cannarsa, C. Sinestrari: An infinite dimensional time optimal control problem, in: Optimization Methods in Partial Differential Equations, S. Cox, I. Lasiecka (eds.), Contemp. Math. 209, Providence (1997) 29–41.

[10] F. H. Clarke: Optimization and Nonsmooth Analysis, Wiley, New York (1983).

[11] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, Nonsmooth Analysis and Control Theory, Springer, New York (1998).

[12] F. H. Clarke, R. J. Stern, P. R. Wolenski: Proximal smoothness and the lower C^2 property, J. Convex Anal. 2 (1995) 117–144.

[13] G. Colombo, V. Goncharov: Variational inequalities and regularity properties of closed sets in Hilbert spaces, J. Convex Anal. 8 (2001) 197–221.

[14] G. Colombo, P. R. Wolenski: The subgradient formula for the minimal time function in the case of constant dynamics in Hilbert space, Journal of Global Optimization, in print.

[15] I. Ekeland, R. Témam: Convex Analysis and Variational Problems, Classics in Applied Mathematics 28, SIAM, Philadelphia (1999).

[16] R. A. Poliquin, R. T. Rockafellar, L. Thibault: Local differentiability of distance functions, Trans. Amer. Math. Soc. 352 (2000) 5231–5249.

[17] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).

[18] P. Soravia: Discontinuous viscosity solutions to Dirichlet problems for Hamilton-Jacobi equations with convex Hamiltonians, Comm. Partial Differ. Equations 18 (1993) 1493–1514.

[19] P. Soravia: Generalized motion of a front propagating along its normal direction, Nonlinear Anal., TMA 22 (1994) 1247–1262.

[20] V. Veliov: Lipschitz continuity of the value function in optimal control, J. Optimization Theory Appl. 94 (1997) 335–363.

[21] P. R. Wolenski, Z. Yu: Proximal analysis and the minimal time function, SIAM J. Control Optim. 36 (1998) 1048–1072.