Journal of Global Optimization **28:** 269–282, 2004.

The subgradient formula for the minimal time function in the case of constant dynamics in Hilbert space

GIOVANNI COLOMBO[1] and PETER R. WOLENSKI[2]
[1]*Dipartimento di Matematica Pura e Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy; (e-mail: colombo@math.unipd.it);* [2]*Department of Mathematics, Louisiana State University, 326 Lockett Hall, Baton Rouge, Louisiana 70803-4918, USA; (e-mail: wolenski@math.lsu.edu)*

(Received and accepted 12 March 2003)

Abstract. The minimal time function with constant dynamics is studied in the context of a Hilbert space. A general formula for the subgradient is proven, and assumptions are identified in which the minimal time function is lower C^2.

Key words: Distance functions, Minimal time function, Subgradient formulas, Variational analysis

1. Introduction

Let X be a real Hilbert space, $S \subseteq X$ be closed, and $F \subseteq X$ be closed, convex, bounded, and with $0 \in \operatorname{int} F$. The results in this note characterize the proximal subgradient of the value function $T_S^F(\cdot): X \to I\!R$ given by

$$T_S^F(x) := \min_{t \geqslant 0} \{t : S \cap \{x + tF\} \neq \emptyset\}. \tag{1}$$

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

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

Our approach is to study $T_S^F(\cdot)$ in analogy to $d_S(\cdot)$, and to pinpoint and analyze the nature of nondifferentiability that arises near S in terms of the so-called F-projection onto S, the latter being the set of points $S \cap \{x + T_S^F(x)F\}$.

We quickly review some recent work on minimal time problems related to our results here. The minimal time function was first characterized as a solution to a Hamilton–Jacobi equation by Bardi [1] using viscosity methods. Soravia [10] extended these results to allow for noncontrollability and more general boundary conditions; see also [13] for a different approach based on invariance. Convexity properties of the minimal time function were derived in [3], where an emphasis was placed on *global* semiconcavity results; assumptions imposed there essentially rule out "semiconvex" behavior which is our focus here. Global semiconvexity was proven in [3], however, in the case of the target being convex and the dynamics linear, and an analogy between the distance function and the minimal time function was also drawn, an approach we are adopting here. Lipschitz estimates are given also in Veliov [12]. Differentiability of minimal time functions related to properties of the velocity set are contained in [2], where the target is taken as a single point.

The underlying state space in all of the aforementioned papers is $I\!R^n$, whereas differentiability properties of the distance function were characterized [5] in a Hilbert space. See also [8] for some additional refinements and localized results, and [6] for related applications with an emphasis on the distinction between finite and infinite dimensional spaces. The aim of the present paper is to extend some of these results from the usual distance function to the case of a general F in the context of a Hilbert space.

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 subgradient of $T_S^F(\cdot)$ in terms of normal vectors to its level sets. A similar propagation formula for general nonlinear systems was proven by Soravia [11] in finite dimensions using contingent cone concepts (Fréchet subgradients and normals); a proximal version in finite dimensions is contained in [13], however the Hilbert space version given here is new. Section 4 contains a formula for the subgradient when S is convex, and Section 5 contains some nonconvex results, and in particular, sufficient conditions for the lower C^2 property in a neighborhood of S.

2. Preliminaries

This section reviews some of the concepts and basic tools to be used in the sequel. See [4] for a fuller development of nonsmooth analysis based on the proximal concepts, and [9] for a somewhat different but exhaustive treatment in finite dimensions.

2.1. BACKGROUND IN VARIATIONAL ANALYSIS

Suppose $S \subseteq X$ is closed and $s \in S$. The *proximal normal cone* $N_S(s)$ to S at s is the set of all $\zeta \in X$ for which there exist $\sigma > 0$ such that

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

If S is convex, then the proximal normal cone coincides with the normal cone of convex analysis. In this case, there is no loss in generality by setting $\sigma = 0$.

The corresponding function concept is the proximal subgradient, defined as follows. Suppose $f: X \to (-\infty, \infty]$ is lower semicontinuous and proper, and let epi $f := \{(x, \alpha) \in X \times I\!R : \alpha \geqslant f(x)\}$ denote the epigraph of f. For $x \in \operatorname{dom} f := \{x \in X : f(x) < \infty\}$, the proximal subgradient $\partial f(x)$ is (the possibly empty) subset of X defined as those ζ satisfying $(\zeta, -1) \in N_{\operatorname{epi} f}\big(x, f(x)\big)$. A more user-friendly description of the proximal subgradient is given by (see [4])

$$\partial f(x) =$$

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

If f is convex, then $\partial f(x)$ coincides with the subgradient of convex analysis. In this case, the above description reduces to taking $\sigma = 0$ and $\eta = \infty$.

2.2. GAUGE FUNCTIONS AND POLARS

We assume throughout that a given set $F \subset X$ is closed, convex, bounded, and with $0 \in \operatorname{int} F$. Recall that the (Minkowski) *gauge* function $\rho_F : X \to [0, \infty]$ associated to F is defined by

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

and the polar F° of F is the set

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

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 $x \in X$ and $r \geqslant 0$) and subadditive ($\rho_F(x+y) \leqslant \rho_F(x) + \rho_F(y)$ 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 $\rho_F(x) \leqslant 1$. Furthermore, the boundedness of F is equivalent to $0 \in \operatorname{int} F^\circ$, and hence F° satisfies the same assumptions being imposed on F. Some further properties are included in the following proposition.

PROPOSITION 2.1.

(*a*) *$v \in$ bdry F if and only if $\rho_F(v) = 1$,*
(*b*) *For all $\zeta \neq 0$ in X,*

$$0 < \rho_{F^\circ}(\zeta) = \max_{v \in F} \langle \zeta, v\rangle < \infty.$$

(*c*) *Define* $\|F\| := \max\{\|v\| : v \in F\}$, *and similarly* $\|F^\circ\| := \max\{\|\zeta\| : \zeta \in F^\circ\}$. *Then for all* $z \in X$,

$$\frac{\rho_F(z)}{\|F^\circ\|} \leqslant \|z\| \leqslant \|F\| \rho_F(z). \tag{2}$$

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

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

Therefore $v + \varepsilon \overline{I\!B} \subseteq F$, and so $v \notin \operatorname{bdry} F$, and (a) is proven.
(b) Let $\zeta \neq 0$. Now $\rho_{F^\circ}(\zeta) > 0$ since $0 \in \operatorname{int} F^\circ$, and is finite since F is bounded. The proof of (b) follows from the calculation

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

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

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

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

Note (b) says that the Hamiltonian as defined in optimal control (see [4]) is $\rho_{F^\circ}(\cdot)$. The next proposition identifies the subgradient of $\rho_F(\cdot)$; the proof is an exercise in convex analysis, see [7].

PROPOSITION 2.2. *Suppose* $v \in X$. *Then*

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

2.3. THE MINIMAL TIME FUNCTION

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

$$T_S^F(x) = \min_{s \in S} \rho_F(s - x). \tag{3}$$

Note that the assumption $0 \in \operatorname{int} F$ implies $T_S^F(x) < \infty$ for all $x \in X$, but the next proposition says in fact $T_S^F(\cdot)$ is Lipschitz. From the optimal control point of view, this is to be expected since the Petrov-type condition is known to characterize the Lipschitz property for nonlinear finite-dimensional systems (see [3, 12, 13]), and $0 \in \operatorname{int} F$ trivially implies the Petrov condition. The proof is straightforward, and therefore omitted.

THEOREM 2.3. *The minimal time function $T_S^F(\cdot)$ is globally Lipschitz on X of rank $\|F^\circ\|$.*

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

$$S(r)=\{y\in X: T_S^F(y)\leqslant r\}=\{y\in X:\{y+rF\}\cap S\neq\emptyset\}. \tag{4}$$

The following theorem contains the special versions of the so-called principle of optimality that are pertinent here.

THEOREM 2.4. (Principle of Optimality). *Suppose $x\notin S$.*

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

$$T_S^F(x-tv)\leqslant T_S^F(x)+t.$$

(*b*) *Let $S(r)$ be as in (4) with $0\leqslant r\leqslant T_S^F(x)$. Then*

$$T_S^F(x)\leqslant r+\min_{z\in S(r)}\rho_F(z-x).$$

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

$$T_S^F(x-tv)\leqslant\rho_F(s-x+tv)\leqslant\rho_F(s-x)+t\rho_F(v)<T_S^F(x)+t+\varepsilon.$$

Letting $\varepsilon\downarrow 0$ proves (a).
(b). Let $\varepsilon > 0$ and suppose $0 \leqslant r\leqslant T_S^F(x)$. There exist $z \in S(r)$ and $s\in S$ so that

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

Therefore

$$T_S^F(x)\leqslant\rho_F(s-z)+\rho_F(z-x)\leqslant r+\min_{z'\in S(r)}\rho_F(z'-x)+2\varepsilon,$$

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

A variant of part (b) above is contained in the following corollary, which although of some interest in itself, is needed in the general infinite dimensional setting where the F-projection set $S\cap\{x+T_S^F(x)F\}$ may be empty. If the latter is nonempty, there are nonetheless always points $s\in S$ that are suboptimal for the optimization problem (3) in the sense of (5) below. The content of the corollary is that from a fixed $x \notin S$, any point z defined on a line originating from x through the associated supoptimal velocity is suboptimal with the same error in the problem of minimizing the F-distance from x to the level set containing z. The case $\varepsilon = 0$ is also covered, and is the case where the suboptimal point is actually optimal.

COROLLARY 1. (A Principle of Suboptimality). *Suppose $x \notin S$, $\varepsilon \geqslant 0$, and $s \in S$ satisfy*

$$\rho_F(s-x) \leqslant T_S^F(x) + \varepsilon. \tag{5}$$

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

$$\bar{t} \leqslant \min_{z \in S(r)} \rho_F(z-x) + \varepsilon. \tag{6}$$

Proof. We have

$$\begin{aligned} r &= \min_{s' \in S} \rho_F(s' - z_{\bar{t}}) \leqslant \rho_F(s - z_{\bar{t}}) \\ &= \rho_F\left(s - x - \bar{t}\frac{s-x}{\rho_F(s-x)}\right) = \rho_F\left(\left[\rho_F(s-x) - \bar{t}\right]\frac{s-x}{\rho_F(s-x)}\right) \\ &= \inf_{t' \geqslant 0}\left\{t' : \frac{\rho_F(s-x) - \bar{t}}{t'}\frac{s-x}{\rho_F(s-x)}\right\} \leqslant \rho_F(s-x) - \bar{t}. \end{aligned}$$

Thus by (5) we have $\bar{t} \leqslant \rho_F(s-x) - r \leqslant T_S^F(x) - r + \varepsilon$, and the final conclusion (6) follows from the previous Theorem, part (b), since it says $T_S^F(x) - r \leqslant \min_{z \in S(r)} \rho_F(z-x)$. □

3. General Formula for $\partial T_S^F(\cdot)$

The result in this section characterizes the proximal subgradient of $T_S^F(\cdot)$ in general terms. The formula has 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.

THEOREM 3.1. *Suppose S is closed and $F \subset X$ is closed, convex, bounded, and with $0 \in \operatorname{int} F$. Suppose $x \notin S$ and $T_S^F(x) = r$. Then*

$$\partial T_S^F(x) = N_{S(r)}(x) \cap \left\{\zeta : \rho_{F^\circ}(-\zeta) = 1\right\},$$

where $S(r)$ is as in (4).

Proof. ($\subseteq$) Let $\zeta \in \partial T_S^F(x)$, and so there exist positive constants η and σ so that

$$T_S^F(y) \geqslant r + \langle \zeta, y - x \rangle - \sigma \|y - x\|^2 \quad \forall y \in x + \eta \overline{I\!B}. \tag{7}$$

If $y \in S(r)$, then $T_S^F(y) \leqslant r$, and it immediately follows from (7) that

$$\langle \zeta, y - x \rangle \leqslant \sigma \|y - x\|^2,$$

or that $\zeta \in N_S^P(r)$.

We next show $\rho_{F^\circ}(-\zeta) \leqslant 1$. Let $v \in F$, and note by the principle of optimality that $y := x - tv$ satisfies $T(y) \leqslant r + t$ for $t \geqslant 0$. Substituting into (7) gives

$$t + r \geqslant T_S^F(y) \geqslant r + \langle \zeta, x - tv - x \rangle - \|x - tv - x\|^2 = r + t\langle \zeta, -v \rangle - t^2 \|v\|^2.$$

Now divide by $t > 0$ and let $t \searrow 0$. Since $v \in F$ is arbitrary, we conclude

$$\rho_{F^\circ}(-\zeta) = \max_{v \in F} \langle -\zeta, v \rangle \leqslant 1. \tag{8}$$

Finally, we show there exists $\bar{v} \in F$ with $\langle \zeta, \bar{v} \rangle \leqslant -1$, which along with (8) implies $\rho_{F^\circ}(-\zeta) = 1$ as desired. For $t > 0$, let $\bar{s}_t \in S$ be so that $\rho_F(s_t - x) \leqslant r + t^2$, 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

$$\frac{1}{t'}\big[s_i - y_i\big] = \frac{\rho_F(s_i - x) - t_i}{t'} \left[\frac{s_i - x}{\rho_F(s_i - x)}\right],$$

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

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

for all i. Clearly y_i is within η of x for large i, and so the previous estimate can be used in conjuction with (7) to obtain

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

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

$$\langle \zeta, \bar{v} \rangle \leqslant -1. \tag{9}$$

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

($\supseteq$) Now suppose $\rho_{F^\circ}(-\zeta) = 1$ and there exists $\sigma' > 0$ so that

$$\langle \zeta, z - x \rangle \leqslant \sigma' \|z - x\|^2 \quad \forall z \in S(r). \tag{10}$$

We must show (7) holds for some choice of η and σ. There are three possibilities for a point y, which we shall consider separately: (i) $T_S^F(y)=r$, (ii) $T_S^F(y)>r$, and (iii) $T_S^F(y)<r$.

(i) The case $T_S^F(y)=r$ is trivial, since (7) then reduces to (10) with $\sigma=\sigma'$.

(ii) Suppose $T_S^F(y)>r$ and $\|y-x\|\leqslant\eta\leqslant 1$. There exists $s\in S$ be such that

$$\rho_F(s-y)<T_S^F(y)+\|y-x\|^2. \tag{11}$$

Set $v:=\frac{s-y}{\rho_F(s-y)}$, and choose $\bar t$ so that $z_t:=y+tv$ satisfies $T_S^F(z_{\bar t})=r$. Since $t_0=0$ satisfies $T_S^F(z_{t_0})=T_S^F(y)>r$ and $t_1=\rho_F(s-y)$ satisfies $T_S^F(z_{t_1})=0$, and therefore such a $\bar t$ exists by the intermediate value theorem. We claim that

$$r+\bar t\leqslant\rho_F(s-y). \tag{12}$$

Indeed,

$$\frac{s-z_{\bar t}}{t'}=\frac{1}{t'}\left[s-y-\bar t\frac{s-y}{\rho_F(s-y)}\right]=\frac{\rho_F(s-y)-\bar t}{t'}\left[\frac{s-y}{\rho_F(s-y)}\right],$$

and so $\frac{s-z_{\bar t}}{t'}\in F$ whenever $t'=\rho_F(s-y)-\bar t$. It follows that

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

which implies (12).

We now combine (11) and (12) to begin the following estimate

$$\begin{aligned}
T_S^F(y)+\|y-x\|^2 &> r+\bar t\\
&\geqslant r+\bar t+\langle\zeta,z_{\bar t}-x\rangle-\sigma'\|z_{\bar t}-x\|^2 \qquad (13)\\
&= r+\bar t+\bar t\langle\zeta,v\rangle+\langle\zeta,y-x\rangle-\sigma'\|z_{\bar t}-x\|^2\\
&\geqslant r+\langle\zeta,y-x\rangle-\sigma'\|z_{\bar t}-x\|^2, \qquad (14)
\end{aligned}$$

where we used (10) to deduce the inequality in (13) (which is valid for $z=z_{\bar t}\in S(r)$), and the assumption $\rho_{F^\circ}(-\zeta)=1$ (which implies $\langle\zeta,v\rangle\geqslant -1$) to deduce the inequality in (14). It is only left to show that if σ is chosen large enough, then $z_{\bar t}$ and σ' in (14) can be replaced by y and σ. Therefore we are done by verifying that

$$\|z_{\bar t}-x\|\leqslant k\|y-x\| \tag{15}$$

holds for some constant k that is independent of y.

The key to proving (15) involves an estimate of $\bar t$ in terms of $\|y-x\|$. The Principle of Suboptimality (Corollary 1, with $\varepsilon=\|y-x\|^2$) implies that

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

and thus

$$\|z_{\bar t}-x\| \leqslant \|z_{\bar t}-y\|+\|y-x\|=\bar t\|v\|+\|y-x\|$$
$$\leqslant \big\{(\|F^\circ\|+1)\|F\|+1\big\}\|y-x\| =: k\|y-x\|.$$

Combining this estimate with (14) and (15) yields

$$T_S^F(y) \geqslant r+\langle\zeta, y-x\rangle-\sigma\|y-x\|^2$$

with $\sigma := \sigma' k+1$, and finishes the proof of (ii).

(iii) Assume now that $T_S^F(y)<r$ and $y\in x+\eta\bar{B}$, where $\eta:=\min\left\{1, \frac{1}{4\sigma'\|F\|}, \frac{1}{16\|F\|^2(\sigma'+\|\zeta\|}\right\}$. Let $\bar v\in F$ be such that

$$\langle\zeta,\bar v\rangle=-1, \tag{16}$$

and let $z_t := y-t\bar v$. We claim that there exists $\bar t\geqslant 0$ so that

$$T_S^F(z_{\bar t})=r \quad \text{with} \quad \bar t\leqslant k\|y-x\| \tag{17}$$

for some constant k independent of y. To prove (17), note that $z_t\in S(r)$ for small t, and so by the intermediate value theorem, $\bar t$ exists with $T_S^F(z_{\bar t})=r$ provided there are t values with $z_t\notin S(r)$. Now z_t does not belong to $S(r)$ if the proximal inequality (10) is violated for $z=z_t$, and we calculate

$$\begin{aligned}
&\sigma'\|z_t-x\|^2-\langle\zeta, z_t-x\rangle=\\
&=\sigma t^2\|\bar v\|^2+2\sigma' t\langle y-x,\bar v\rangle+\sigma'\|y-x\|^2-\langle\zeta,y-x\rangle+t\langle\zeta,\bar v\rangle\\
&=\sigma'\|\bar v\|^2t^2+\big\{2\sigma'\langle y-x,\bar v\rangle-1\big\}t+\big\{\sigma\|y-x\|^2-\langle\zeta,y-x\rangle\big\}\\
&=: at^2+bt+c,
\end{aligned}$$

where we have used (16). Note that $b\leqslant \frac{-1}{2}$ since $2\sigma'\langle y-x,\bar v\rangle\leqslant 2\sigma\|F\|\,\|y-x\|$, and that the above quadratic function in t has real roots since $4ac\leqslant 4\|F\|^2(\sigma'+\|\zeta\|)\|y-x\|<\frac{1}{4}\leqslant b^2$. Its smallest root is given by

$$\hat t := \frac{-b-\sqrt{b^2-4ac}}{2a}=\frac{4ac}{2a[-b+\sqrt{b^2-4ac}]}. \tag{18}$$

It follows that $z:=z_t$ for t slightly larger than $\hat t$ is such that (10) is violated, and consequently $z_t\notin S(r)$ for those t. Hence there exists $\bar t\leqslant\hat t$ so that $T_S^F(z_{\bar t})=r$. We obtain the estimate in (17) of $\bar t$ from (18): Since $\|x-y\|\leqslant\eta$, we have $-b = 1-2\sigma\langle y-x,\bar v\rangle \geqslant 1-2\sigma\|\bar v\|\eta\geqslant\frac{1}{2}$. Thus it follows from (18) that $\bar t\leqslant\hat t\leqslant 4c\leqslant 4\{\sigma+\|\zeta\|\}\|y-x\| =: k\|y-x\|$, and thus (17) is valid as claimed.

We now return to proving (7). We have by the principle of optimality again that

$$T_S^F(z_{\bar t})\leqslant T_S^F(y)+\bar t,$$

from which

$$T_S^F(y) \geqslant T_S^F(z_{\bar t}) - \bar t = r - \langle \zeta, -\bar t \bar v\rangle$$
$$= r + \langle \zeta, y-x\rangle - \langle \zeta, y - \bar t \bar v - x\rangle = r + \langle \zeta, y-x\rangle - \langle \zeta, z_{\bar t} - x\rangle \tag{19}$$

follows from (16). Now $z_{\bar t} \in S(r)$ by (17), so by (10),

$$\langle \zeta, z_{\bar t} - x\rangle \leqslant \sigma' \|z_{\bar t} - x\|^2. \tag{20}$$

Also, from the $\bar t$-estimate in (17), we obtain

$$\|z_{\bar t} - x\| \leqslant \|y-x\| + \bar t \|\bar v\| \leqslant \big[1 + k\|F\|\big]\|y-x\|. \tag{21}$$

Substituting (20) and (21) into (19) yields (7) with $\sigma = \sigma'(1+k\|F\|)^2$. □

The following result is a corollary to the previous proof (see (9)), and will be used below.

COROLLARY 2. *Suppose $x \notin S$, $s \in \Pi_S^F(x)$, and $-\zeta \in \partial \rho_F(s-x)$. Then*

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

4. The Case Where *S* is Convex

The case where S is convex can be completely described in a global manner. 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*

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

THEOREM 4.2. *Suppose S is convex. Then*

(*a*) *$T_S^F(\cdot)$ is convex on X;*
(*b*) *For each $x \in X$, 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*
(*d*) *The convex subgradient $\partial T_S^F(\bar x)$ is given by*

$$\partial T_S^F(x) = \left\{ \frac{\zeta}{\rho_{F^\circ}(-\zeta)} : \exists \bar s \in \Pi_S^F(\bar x)\ with\ \zeta \in SEP(S/F, \bar s, \bar x)\right\}$$
$$= \left\{ \frac{\zeta}{\rho_{F^\circ}(-\zeta)} : \forall \bar s \in \Pi_S^F(\bar x)\ with\ \zeta \in SEP(S/F, \bar s, \bar x)\right\}$$

The proof is omitted here due to lack of space, but details will be provided elsewhere. The finite-dimensional version of the theorem can be derived from [9], Theorem 10.13, page 433.

COROLLARY 3. *Suppose S is convex and* $x \notin S$. *Then*

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

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

5. Results for Nonconvex *S*

We now consider conditions on F and nonconvex S for which $T_S^F(\cdot)$ has some regularity properties in a neighborhood of S. The following theorem says that one inclusion of the equality in Corollary 3 holds for general closed sets S.

THEOREM 5.1. *Suppose* $S \subseteq X$ *is closed,* $x \notin S$, *and* $\Pi_S^F(x) \neq \emptyset$. *Then the following inclusion holds*:

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

Proof. The inclusion is trivial if $\partial T_S^F(x) = \emptyset$, so suppose $\zeta \in \partial T_S^F(x)$, and set $r := T_S^F(x)$. Recall Theorem 3.1, which says $\rho_{F^\circ}(-\zeta) = 1$ and $\zeta \in N_{S(r)}(x)$. By Proposition 2.2, it suffices to show

$$\zeta \in \left[-N_F\left(\frac{s-x}{\rho_F(s-x)}\right)\right] \cap N_S(s) \quad \forall s \in \Pi_S^F(x). \tag{22}$$

Let $s \in \Pi_S^F(x)$, and set $\bar{v} := \frac{s-x}{\rho_F(s-x)}$. First note by Corollary 2 that $\langle -\zeta, \bar{v} \rangle = 1$. Also, since $\rho_{F^\circ}(-\zeta) = 1$, we have $\langle -\zeta, v \rangle \leqslant 1$ for all $v \in F$. Hence $\langle -\zeta, v - \bar{v} \rangle \leqslant 0 \ \ \forall v \in F$, and so $\zeta \in -N_F(\bar{v})$.

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

$$\langle \zeta, y - x \rangle < \sigma \|y - x\|^2 \quad \forall y \in S(r). \tag{23}$$

Let $s' \in S$, and note that $y := s' + x - s$ belongs to $S(r)$ (since $T_S^F(y) \leqslant \rho_F(s' - y) = \rho_F(s-x) = r$). Since $s' - y = s - x$ and $y - x = s' - s$, we have by (23) that

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

Hence $\zeta \in N_S(s)$. □

Our final goal is to identify hypotheses so that the opposite inclusion in Theorem 5.1 holds, and thereby obtain regularity results for $T_S^F(\cdot)$. This is well-understood for the case of $F=\overline{I\!B}$, see [5], [8], [6]. A variety of equivalent conditions were shown in [5] for the distance function to be C^1 in a neighborhood of S, and such sets were labeled *proximally smooth*. By Theorem 5.1, it is clear that if one seeks $\partial T_S^F(x)\neq\emptyset$ for x in a neighborhood of S, then S must have plentiful proximal normal vectors, and proximal smoothness is precisely this. We use the following as the definition of proximal smoothness: there exists $\sigma>0$ so that for all $s\in S$ and $\zeta\in N_S(s)$,

$$\langle\zeta, s'-s\rangle\leqslant\sigma\|s'-s\|^2 \quad \forall s'\in S. \tag{24}$$

Notice that σ is independent of $s\in S$. In addition to proximal smoothness, the following theorem hypothesizes a sort of one-sided Lipschitz condition of the F-projection map, and just as in the convex case, concludes that (22) holds as an equality. However, it is still not clear if (25) always holds for S convex, although it can be shown in some cases the projection map is singleton-valued and Lipschitz, which then immediately implies (25).

THEOREM 5.2. *Suppose $x\notin S$ is such that for all there exist constants $\eta>0$, $k>0$, so that*

$$\Pi_S^F(y)\subseteq\Pi_S^F(x)+k\|y-x\|\overline{I\!B} \quad \forall y\in x+\eta\overline{I\!B}, \tag{25}$$

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

$$\partial T_S^F(x)=-\partial\rho_F(s-x)\cap N_S(s)$$

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

Proof. The inclusion "$\subseteq$" is the result of Theorem 5.1. In view of Theorem 3.1 and Proposition 2.2, the opposite inclusion "$\supseteq$" follows if it can be shown that

$$\left[-N_F\left(\frac{s-x}{\rho_F(s-x)}\right)\right]\cap N_S(s)\subseteq N_{S(r)}(x) \tag{26}$$

for all $s\in\Pi_S^F(x)$, where $r:=T_S^F(x)$. Note the left side of (26) is independent of the particular $s\in\Pi_S^F(x)$ since $-\partial\rho_F(s-x)\cap N_S(s)$ is and the fact that these sets differ only by scaling (Proposition 2.2). Thus it suffices to show there exists $s\in\Pi_S^F(x)$ so that (26) holds.

There exist constants $k>0$ and $\eta>0$, so that $y\in x+\eta\overline{I\!B}$ implies (25) holds, and there exists $\sigma>0$ so that (24) holds for all $s\in S$. Suppose ζ belongs to the left side of (25) for some (and therefore all) $s\in\Pi_S^F(x)$. Now let $y\in x+\eta\overline{I\!B}\cap N_{S(r)}(x)$, and select any $s'\in\Pi_S^F(y)$. By (24), there exists $s\in\Pi_S^F(x)$ so that

$$\|s'-s\|\leqslant k\|y-x\|. \tag{27}$$

Now we write

$$\langle \zeta, y-x\rangle = r\left\langle -\zeta, \frac{s'-y}{r} - \frac{s-x}{r}\right\rangle + \langle \zeta, s'-s\rangle, \tag{28}$$

and note that $\rho_F(s'-y) = T_S^F(y) \leqslant r$ implies $\rho_F\left(\frac{s'-y}{r}\right) \leqslant 1$, or that $\frac{s'-y}{r}$ belongs to F. Since $-\zeta \in N_F\left(\frac{s-x}{r}\right)$, the first term on the righthand side of (28) is thus nonpositive. The second term is bounded by $\sigma\|s'-s\|^2$ by (24), and so by (27) and (28), we have

$$\langle \zeta, y-x\rangle \leqslant \sigma k^2 \|y-x\|^2.$$

This says $\zeta \in N_{S(r)}(x)$ and finishes the proof of (26), and consequently of the theorem. □

We conclude with a corollary concerning a sufficient condition for semiconvexity (which is called lower C^2 in [9], [5]) of $T_S^F(\cdot)$ near S.

COROLLARY 4. *Suppose S is proximally smooth, the F-projection $\Pi_S^F(x)$ of each point x in an open neighborhood U of S is unique, and $x \mapsto \Pi_S^F(x)$ is Lipschitz on U. Then $\partial T_S^F(x) \neq \emptyset$ for all $x \in U$ and $T_S^F(\cdot)$ is lower C^2 on U.*

Acknowledgements

This work was partially supported by GNAMPA and PRIN "Analisi e controllo di equazioni di evoluzione deterministiche e stocastiche." P.R.W.'s work was partially supported by NSF grant DMS-9972241.

References

1. Bardi, M. (1989), A boundary value problem for the minimal time problem, *SIAM J. Control and Optimization* 27, 776–785.
2. Bressan, A. (1980), On two conjectures by Hájek, *Funkc. Ekv.* 23, 221–227.
3. Cannarsa, P. and Sinestrari, C, (1995), Convexity properties of the minimum time function, *Calculus of Variations* 3, 273–298.
4. Clarke, F.H., Ledyaev, Yu.S., Stern, R.J. and Wolenski, P.R. (1998), *Nonsmooth Analysis and Control Theory*, Springer, New York.
5. Clarke, F.H., Stern, R.J., and Wolenski, P.R. (1995), Proximal smoothness and the lower C^2 property, *Journal of Convex Analysis* 2, 117–144.
6. Colombo, G. and Goncharov, V. (2001), Variational inequalities and regularity properties of closed sets in Hilbert spaces, *Journal of Convex Analysis* 8, 197–221.
7. Ekeland, I. and Témam, R. (1999), *Convex Analysis and Variational Problems*, Classics in Applied Mathematics, SIAM, Philadelphia.
8. Poliquin, R.A., Rockafellar, R.T. and Thibault, L. (2000),Local differentiability of distance functions, *Transactions of the American Mathematical Society* 352, 5231–5249.
9. Rockafellar, R.T. and Wets, R.J-B. (1998), *Variational Analysis*, Springer, Berlin.

10. Soravia, P. (1993), Discontinuous viscosity solutions to Dirichlet problems for Hamilton–Jacobi equations with convex Hamiltonians, *Communications Partial Differential Equations* 18, 1493–1514.
11. Soravia, P. (1994), Generalized motion of a front propagating along its normal direction, *Nonlinear Analysis TMA* 22, 1247–1262.
12. Veliov, V. (1997), Lipschitz continuity of the value function in optimal control, *Journal of Optimization Theory* 94, 335–363.
13. Wolenski, P.R. and Yu, Z. (1998), Proximal analysis and the minimal time function, *SIAM J. Control Optim.* 36, 1048–1072.