To appear in
Dynamics of Continuous, Discrete and Impulsive Systems
http:monotone.uwaterloo.ca/~journal

SEMICONCAVITY OF THE MINIMUM TIME FUNCTION FOR DIFFERENTIAL INCLUSIONS

Piermarco Cannarsa[1] , Francesco Marino[2] and Peter Wolenski[3]

[1]Department of Mathematics
Tor Vergata University of Rome, Rome, Italy

[2]Department of Mathematics
Tor Vergata University of Rome, Rome, Italy

[3]Department of Mathematics
Louisiana State University, Baton Rouge, Louisiana 70803-4918, USA
Corresponding author email:cannarsa@mat.uniroma2.it(P.Cannarsa),
wolenski@math.lsu.edu(P.Wolenski)

Abstract. In this paper we consider the Minimum Time Problem with dynamics given by a differential inclusion. We prove that the minimum time function is semiconcave under suitable hypotheses on the multifunction F.

Keywords. Differential Inclusion, Semiconcavity, Optimality, Minimum Time Function.

AMS (MOS) subject classification: This is optional. But please supply them whenever possible.

1 Introduction

The minimum time problem consists of minimizing the time T over all trajectories of a controlled dynamical system that originate from an initial point $x \in \mathbb{R}^n$ and terminate on a compact target set $\mathcal{K} \subseteq \mathbb{R}^n$. The main goal of this paper is to develop and extend aspects of the considerable theory of this subject to situations where the controlled dynamics take the form of a differential inclusion. Specifically, we shall study the problem

$$\min T, \tag{1}$$

where the minimization is over all absolutely continuous arcs $y(\cdot)$ that satisfy the differential inclusion

$$\dot{y}(t) \in F\big(y(t)\big) \quad \text{a.e. } t \in [0, T] \tag{2}$$

and the initial and terminal conditions

$$y(t) = x \quad \text{and} \quad y(T) \in \mathcal{K}. \tag{3}$$

Here and throughout the paper, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a given multifunction. As a function of the initial parameter x, the minimum time function $T(x)$ equals the optimal in (1) and is our main object of study.

There is an extensive literature on the minimum time problem, perhaps originating with Hermes and Lasalle [12], but having since grown beyond our capacity here to mention everyone who has contributed to the topic. The relationship between the regularity of the minimum time function and the local controllability property is well known in classical control theory. The first definition of Petrov Condition is due to [14, 13], in which the author investigated the Lipschitz property of T when the target set is a singleton. More general results have been obtained over the last two decades. In [3], under the regularity assumption for the target set, $T(\cdot)$ is shown to be Lipschitz continuous if and only if controllability conditions are satisfied. The Hölder continuity of the minimum time function when the target is a lower-dimensional manifold is studied in [17]. Further regularity properties of the minimum time function for a nonlinear control system can be found in [7, 8] in which under a Petrov controllability assumption, $T(\cdot)$ is shown to be semiconcave if the distance function from a target set is semiconcave. Moreover a condition of Petrov type for a general target is shown to be necessary and sufficient for the Lipschitz continuity of T. A Necessary and sufficient condition for local Lipschitz continuity of the value function is provided in [18] when the target is an arbitrary closed set and the dynamics is given by a differential inclusion with a measurability assumption on t. In [19] a Petrov-type modulus condition is also considered to prove the local continuity of T near the target set. The semiconcavity of a value function for a general exit time problem is given in [16], where suitable smoothness assumptions on the dynamics of the system are required but the target set can be completely general. In [5] structural properties of a nonlinear control system ensuring that the attainable set satisfies a uniform interior sphere condition are detected. As a consequence of this result, a semiconcavity property for the value function with a general target is obtained. See [2] and [8] for further references on time optimal control problems and semiconcave functions.

Our main goal here is to extend the semconcavity result from [6, 7] to a class of differential inclusions that were recently introduced in [9]. In a companion paper to this one, we shall address the other main results from [6, 7], where the dynamic framework contains a control parametrization with C^{1+}-dependence; that is, the gradient in the state variable exists and is locally Lipschitz. However, the conclusions in these results have a nature that depends only on the admissible velocity sets $F(x)$ and not on a given parametrization. Thus it is natural to seek conditions only on the multifunction F that ensure similar conclusions. A full discussion of this issue as it relates to the Mayer problem is presented in [9], and those considerations are pertinent for the minimum time problem that is being addressed in the current paper.

A well-developed theory for differential inclusions exists (see e.g. [10, 11])

under the following collection of Standing Hypotheses:

$$(SH) \quad \begin{cases} \mathbf{1)}\, F(x) \text{ is nonempty, convex, and compact for each } x \in \mathbb{R}^n, \\ \mathbf{2)}\, F \text{ is locally Lipschitz with respect to the Hausdorff metric,} \\ \mathbf{3)}\, \text{there exists } r > 0 \text{ so that } \max\{|v| : v \in F(x)\} \leq r(1+|x|). \end{cases}$$

Recall the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ associated to F is defined by

$$H(x,p) = \sup_{v \in F(x)} \langle v, p \rangle, \tag{4}$$

and that there is a one-to-one correspondence between Hamiltonian functions that are convex and positively homogeneous in p and multifunctions F with closed and convex values. The relationship is

$$v \in F(x) \quad \Longleftrightarrow \quad \langle v, p \rangle \leq H(x,p) \quad \forall p \in \mathbb{R}^n.$$

Assumptions on H are therefore intrinsic to F and do not depend on a particular parameterization. The following additional hypotheses will be invoked, and are stated in terms of Hamiltonian H: For each compact convex subset $K \subseteq \mathbb{R}^n$ and $p \neq 0$ we have

$$(H) \quad \begin{cases} \mathbf{1)} \quad \text{there exists a constant } c \geq 0 \text{ (depending on } K\text{) so that } x \mapsto H(x,p) \\ \qquad \text{is semiconvex on } K \text{ with constant } c|p|, \\ \mathbf{2)} \quad \text{the gradient } \nabla_p H(x,p) \text{ exists and is Lipschitz in } x \text{ on } K, \text{ and} \\ \qquad \text{uniformly so over } p \text{ in compact subsets of } \mathbb{R}^n \setminus \{0\}. \end{cases}$$

The definition of semiconvexity and other preliminaries will be recalled in the next section. A method to generate examples of multifunctions satisfying (SH) and (H) is given in [9].

The main proof technique behind the semiconcavity results in [4, 6, 8] and other earlier work relied on a priori estimates of solutions to smoothly parameterized differential equations. Our proof technique here relies heavily on the maximal principle. The main contribution of this paper and its companion is to show how the broad theory of minimum time optimal control can be carried out when smooth parameterization assumptions are replaced by the assumption (H).

We provide background material and preliminary results in Section 2, and our main semiconcavity result is given and proved in Section 3.

2 Preliminaries

We review the basic concepts from nonsmooth analysis used in the sequel. Standard references are [8, 11].

2.1 General considerations and semiconcavity

We denote by $\mathbb{B}$ the unit ball of $\mathbb{R}^n$ centered at 0, and by $\mathbb{B}_r(x) := x + r\mathbb{B}$ the ball centered at x of radius r. The closed balls are denoted similarly but replace $\mathbb{B}$ by $\overline{\mathbb{B}}$. Hereafter, $\overline{S}$ and $\text{int}(S)$ stand for the closure and interior, respectively, of any given set $S \subset \mathbb{R}^n$.

Let $K \subseteq \mathbb{R}^n$ be a closed set and $x \in K$. A vector $\nu \in \mathbb{R}^n$ is a *proximal normal* to K at x provided there exists $\sigma > 0$ so that $\langle \nu, y - x\rangle \leq \sigma \|y - x\|^2$ for all $y \in K$. The *proximal normal cone* to K at x consists of all the proximal normals to K at x and is denoted by $\mathcal{N}_K^P(x)$. One has $\nu \in \mathcal{N}_K^P(x)$, $\|\nu\| = 1$ if and only if there exists $r > 0$ so that $\overline{\mathbb{B}}_r(x + r\nu) \cap K = \{x\}$. The Clarke normal cone $\mathcal{N}_K(x)$ is defined as

$$\overline{\text{co}}\big\{\zeta : \exists x_i \to x,\, \zeta_i \to \zeta, \zeta_i \in \mathcal{N}_K^P(x_i)\big\}\,,$$

where $\overline{\text{co}}$ denotes taking the closed convex hull.

Now suppose $f : \mathbb{R}^n \to (-\infty, +\infty]$ is lower semicontinuous and

$$\text{epi } f = \{(x, \alpha) : \alpha \geq f(x)\}$$

is its (closed) epigraph. For $x \in \text{dom } f := \{x : f(x) < \infty\}$, the set of proximal subgradients of f at x is denoted by $\partial_P f(x)$ and equals

$$\big\{\nu \in \mathbb{R}^n : (\nu, -1) \in \mathcal{N}^P_{\text{epi } f}(x, f(x))\big\}\,.$$

The limiting subgradient $\partial_L f(x)$ equals

$$\big\{\nu : \exists x_i \to x,\, \nu_i \to \nu, \nu_i \in \, \partial_P f(x_i)\big\}\,,$$

and the Clarke subgradient $\partial f(x)$ is defined as $\partial f(x) = \overline{\text{co}}\, \partial_L f(x)$. Recall f is locally Lipschitz provided for all compact sets $K \subset \mathbb{R}^n$, there exists a constant k (called the rank of f on K) so that x, $y \in K$ implies $|f(x) - f(y)| < k|x - y|$). For a closed set $K \subseteq \mathbb{R}^n$, the distance function $d_K(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is defined by $d_K(x) := \inf\{\|y - x\| : y \in K\}$, and is globally Lipschitz of rank one and satisfies $\partial d_K(x) = \mathcal{N}_K(x) \cap \overline{\mathbb{B}}$ for all $x \in \text{bdry } K$.

The property of *semiconcavity* has several characterizations for both sets and functions, and the following proposition lists some of the functional properties that will be used in the sequel.

Proposition 1. *For a convex set $K \subseteq \mathbb{R}^n$, a function $f : K \to \mathbb{R}$, and a constant $c \geq 0$, the following properties are equivalent:*

(1) for all x_0, $x_1 \in K$ and $0 \leq \lambda \leq 1$, one has

$$(1 - \lambda)f(x_0) + \lambda f(x_1) - f(x_\lambda) \leq c\lambda(1 - \lambda)|x_1 - x_0|^2,$$

where $x_\lambda = (1 - \lambda)x_0 + \lambda x_1$, *(2) f is continuous and, for all $x \in K$ and $z \in \mathbb{R}^n$ with $x \pm z \in K$, one has*

$$f(x+z) + f(x-z) - 2f(x) \le 2c|z|^2\,,$$

(3) the map $x \mapsto f(x) - c|x|^2$ is concave.

Moreover, if K is open and f is locally Lipschitz, then any of the above properties holds true if and only if

(4) for each $x \in K$ and $\nu \in \partial f(x)$, one has

$$f(y) \le f(x) + \langle \nu, y - x\rangle + c|y-x|^2$$

for all $y \in K$.

See [8] for further details and a thorough development. A Lipschitz function f satisfying property (1) of Proposition 1 is called *(linearly) semiconcave* on K with constant c. The above "concave" concepts have "convex" counterparts by reversing the inequalities and the signs in the quadratic terms. In other words, f is *semiconvex* on K if and only if $-f$ is semiconcave on K.

We note that semiconcavity of f is the same as the property f *being lower* C^2 that is presented in [15, Definition 10.29], where f is given a representation as an infinum of a parameterized family of C^2 functions. Related to that definition is the following simple fact which can be found in [8, Proposition 2.1.5].

Proposition 2. *Suppose $\{f_\alpha(\cdot)\}_{\alpha\in A}$ is a family of semiconcave functions defined on the convex set K with a constant c independent of α belonging to the compact index set A. Then $f(x) = \inf\{f_\alpha(x) : \alpha \in a\}$ is semiconcave on K with constant c.*

Note that the previous result implies that if F is parameterized as $F(x) = \{f(x,u) : u \in U\}$ where f has C^{1+}-dependence in x and continuous in u with U compact, then for each $p \in \mathbb{R}^n$, one has $x \mapsto H(x,p)$ is semiconvex. That is, such an F satisfies our new hypothesis (H1).

The Hausdorff distance between two compact subsets S_1 and S_2 of $\mathbb{R}^n$ is defined as

$$\mathrm{dist}_{\mathcal{H}}(S_1,S_2) = \max\big\{\mathrm{dist}^+_{\mathcal{H}}(S_1,S_2), \mathrm{dist}^+_{\mathcal{H}}(S_2,S_1)\big\},$$

where $\mathrm{dist}^+_{\mathcal{H}}(S,S') = \inf\big\{\varepsilon : S \subseteq S' + \varepsilon\mathbb{B}\big\}$. A multifunction $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with compact values is Lipschitz on $K \subseteq \mathbb{R}^m$ provided there exists a constant k (called a Lipschitz rank of F) so that $\mathrm{dist}_{\mathcal{H}}\big(F(x),F(y)\big) \le k|x-y|$ for all x, $y \in K$. This holds if and only if for each $p \in \mathbb{R}^n$, $x \mapsto H(x,p)$ is a Lipschitz function on K with constant $k|p|$. The mid-point property for a multifunction F on a convex subset K is that

$$\mathrm{dist}^+_{\mathcal{H}}\Big(2F(x), F(x+z) + F(x-z)\Big) \le c|z|^2 \tag{5}$$

holds whenever x, $x \pm z \in K$. This is equivalent to $x \mapsto H(x,p)$ being semiconvex on K for each $p \in \mathbb{R}^n$ with constant $c|p|$, which is the new assumption (H1) we shall impose in proving our main result.

If the above function or multifunction is defined on all of $\mathbb{R}^n$, then all of the above function and multifunction concepts can be quantified as being *local*, by which is meant that the said property holds in every convex compact neighborhood K of each point in $\mathbb{R}^n$ with the constant depending on K.

The first part of the following result comes from [10, Prop 3.2.4(e)], but we give a simpler proof based on the assumption (H1). The second part is a partial inclusion analogous to [10, Prop 3.2.4(d)], but with convexity replaced by semiconvexity.

Proposition 3. *Suppose F satisfies (SH) and $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is as in (4) and satisfies (H1).*

(1) For any $x \in \mathbb{R}^n$ and for any $p \in \mathbb{R}^n$, we have for $\lambda > 0$ that $H(x, \lambda p) = \lambda H(x,p)$ and

$$\partial_x H(x, \lambda p) = \lambda \partial_x H(x,p).$$

(2) For $(u,v) \in \partial H(x,p)$, one has $u \in \partial_x H(x,p)$ and $v \in \partial_p H(x,p)$. That is,

$$\partial H(x,p) \subseteq \partial_x H(x,p) \times \partial_p H(x,p)$$

Proof. The positive homogeneity property $H(x, \lambda p) = \lambda H(x,p)$ for all $\lambda > 0$ holds as seen directly from the definition (4). Suppose $u \in \partial_x H(x, \lambda p)$ with $\lambda > 0$. By (H1), the following inequality holds for any y in a neighborhood of x:

$$H(y, \lambda p) - H(x, \lambda p) \geq \langle u, y - x \rangle - o(y - x)$$

This is equivalent by positive homogeneity to

$$H(y,p) - H(x,p) \geq \left\langle \frac{u}{\lambda}, y - x \right\rangle - o(y - x),$$

which says that $\frac{u}{\lambda} \in \partial_x H(x,p)$; that is

$$u \in \lambda \partial_x H(x,p),$$

which proves (1). The proof of (2) is given in [9, Lemma 2.1] □

2.2 Differential inclusions

There are several books (cf. [10, 1, 11]) that contain the basics of differential inclusion theory. It is well-known that the assumptions of (SH) imply that if $T(x) < +\infty$, then the minimum time problem (1)-(3) has an optimal solution. We next recall the Necessary Conditions for an optimal solution. See [10, Theorem 3.2.6]. Actually recent and more general necessary conditions can be found in the literature, but this one is enough for our aims because

(H1) implies the subgradients of H split into its (x,p) components, which relays the effect of Proposition 3(2).

Theorem 1 (Necessary Conditions)**.** *Assume that (SH) and (H1) hold. Suppose $x(\cdot)$ is an optimal solution for the minimum time problem. Then there exists an absolute continuous arc $p : [0,T] \mapsto \mathbb{R}^n$ such that for a.e. $t \in [0,T]$*

$$\begin{cases} \dot{x}(t) & \in & \partial_p H(x(t), p(t)) \\ -\dot{p}(t) & \in & \partial_x H(x(t), p(t)) \\ -p(T) & \in & \mathcal{N}_{\mathcal{K}}(x(T)) \end{cases} \tag{NC}$$

In addition, $p(\cdot)$ can be normalized to satisfy $|p(T)| = 1$.

The last normalized statement follows from Proposition 3(1).

Remark 1. It is noteworthy that the dual arc p never vanishes. Indeed let $C|p|$ be the Lipschitz constant for $H(\cdot, p)$, then

$$|\dot{p}(t)| \le C|p(t)| \qquad t \in [0,T]$$

Therefore, by Gronwall's lemma, easily we obtain that for any $t \in [0, T(x)]$ either $p(t) \neq 0$ or $p(t) \equiv 0$.

We next turn to an auxiliary ODE that resembles a "feedback map" associated with a particular dual arc. The differentiability statement in (H2) is equivalent to the argmax set of $v \mapsto \langle v, p\rangle$ over $v \in F(x)$ being a singleton $\bar{v}$. We denote this unique element by $F_p(x) = \bar{v}$, which equals $\nabla_p H(x,p)$, and note that $p \mapsto F_p(x)$ is continuous. The Lipschitz statement in (H2) says that $x \mapsto F_p(x)$ is locally Lipschitz. The following result is the key ingredient that allows us to replace smooth parameterizations by (H). We quote it from [9, Proposition 3] and give some important immediate consequences.

Proposition 4. *Assume that (SH) and (H) hold, and $p(\cdot)$ is an absolutely continuous arc defined on $[0,T]$ with $p(t) \neq 0$ for all $t \in [0,T]$. Then for each $x \in \mathbb{R}^n$, consider the initial value problem*

$$\begin{cases} \dot{x}(t) = F_{p(t)}\big(x(t)\big) \qquad a.e.\ t \in [0,T] \\ x(0) = x \end{cases} \tag{IVP}$$

(1) (IVP) has a unique solution $y(\cdot; 0, x)$ that is a solution of (2).

(2) The arc $y(\cdot; 0, x) := x(\cdot)$ satisfies

$$\langle p(t), \dot{x}(t)\rangle = H\big(x(t), p(t)\big) \quad a.e.\ t \in 0, T].$$

(3) As a function of the initial condition, $x \mapsto y(t; 0, x)$ is locally Lipschitz on $\mathbb{R}^n$ with constant independent of $t \in [0,T]$. That is, for each $M > 0$ there exists a constant k so that $|x_i| \le M$, $i = 1, 2$, implies

$$\big|y(t; 0, x_1) - y(t; 0, x_2)\big| \le k|x_1 - x_2|.$$

The following is a direct consequence of (NC).

Proposition 5. *If $x(\cdot)$ is an optimal solution to (1)-(3) and $p(\cdot)$ is its associated adjoint arc, then $x(\cdot)$ solves (IVP).*

2.3 Assumptions on the target $\mathcal{K}$

In addition to assumptions on the dynamic data F, we shall impose two assumptions on the target $\mathcal{K}$. The first is the well-known Petrov condition (see [8, Definition 8.2.2]).

(PC) $\exists\delta > 0$ so that $H(x, -\nu) \geq \delta\|\nu\|, \quad \forall x \in \text{bdry } \mathcal{K},\ \nu \in \mathcal{N}^P_{\mathcal{K}}(x)$.

Assumption (PC) turns out to be equivalent to the minimal time function $T(\cdot)$ being Lipschitz in a neighborhood of $\mathcal{K}$, but is also useful in obtaining a bound of $T(\cdot)$ in terms of the distance function $d_{\mathcal{K}}(x) := \inf\{\|x-y\| : y \in \mathcal{K}\}$. The following can be found at [8, Theorem 8.2.3].

Proposition 6. *Assume F satisfies (SH) and $\mathcal{K} \subseteq \mathbb{R}^n$ is compact. Then (PC) is equivalent to the existence of constants $\rho, m > 0$ so that*

$$T_{\mathcal{K}}(x) \leq m d_{\mathcal{K}}(x) \quad \forall x \in \mathcal{K}_\rho,$$

where $\mathcal{K}_\rho := \{x : d_{\mathcal{K}}(x) \leq \rho\}$.

The second assumption we shall require is known as the Interior Sphere Property (ISP). It has been invoked in similar contexts as we are using here, for example in [7] and [8], and is stated as follows.

(ISP) $\exists r > 0$ so that $\forall x \in \mathcal{K}, \exists y \in \mathcal{K}$ such that $x \in \overline{\mathbb{B}}_r(y) \subseteq \mathcal{K}$.

The constant $r > 0$ in this definition is the size of *realizability* of (ISP), and when this dependence is relevant (as it is in the next result) we write $(\text{ISP})_r$ for this assumption. There is an intrinsic relationship between (ISP) and the regularity properties of the distance function restricted to $\mathbb{R}^n/\mathcal{K}$ (see [8, Proposition 2.2.2] for the proof).

Proposition 7. *If $(\text{ISP})_r$ holds, then the distance function $d_{\mathcal{K}}(\cdot)$ is locally semiconcave on $\mathbb{R}^n \setminus \mathcal{K}$ with constant $\frac{1}{r}$.*

Actually we will use a slightly stronger property than the previous result. Recall that the signed distance from the boundary of a nonempty set $S \subset \mathbb{R}^n$ is given by

$$\bar{d}_S(x) = d_S(x) - d_{\mathbb{R}^n \setminus S}(x) \qquad \forall x \in \mathbb{R}^n\,. \tag{6}$$

Obviously for a ball $\mathbb{B}_r(x_0)$, the signed distance turns out to be

$$\bar{d}_{\mathbb{B}_r(x_0)}(x) = \|x - x_0\| - r \qquad \forall x \in \mathbb{R}^n\,, \tag{7}$$

and it can readily checked that it is semiconcave on each convex subset of $\mathbb{B}_R(0) \cap \mathbb{R}^n \setminus \{\mathbb{B}_{\frac{r}{2}}(x_0)\}$ with a constant C depending only on r and $R > r$. The interesting consequence of the $(ISP)_r$ property that will be useful in the proof of our main result is the following.

Proposition 8. *Suppose $\mathcal{K} \subset \mathbb{R}^n$ is compact and satisfying $(ISP)_r$ for some $r > 0$, and let $R > r$. Then there exists a constant $C = C(r, R) > 0$ such that*

(1)
$$\bar{d}_{\mathcal{K}}(x+z) + \bar{d}_{\mathcal{K}}(x-z) - 2\bar{d}_{\mathcal{K}}(x) \leq C|z|^2 \tag{8}$$
holds for all $x, z \in \mathbb{B}_R(0)$ with $x \in \mathbb{R}^n \setminus \mathrm{int}(K)$.

(2) Let $x \in$ bdry $\mathcal{K}$ and denote by $B := \overline{\mathbb{B}}_r(x_B) \subset \mathcal{K}$ the ball touching the boundary of $\mathcal{K}$ at x. Then
$$\bar{d}_{\mathcal{K}}(y) \leq \langle \nabla \bar{d}_B(x), y - x \rangle + C|y-x|^2 \tag{9}$$
for all $y \in \mathbb{B}_R(0)$.

Proof. Fix $x, z \in \mathbb{B}_R(0)$ such that $x \in \mathbb{R}^n \setminus \mathrm{int}(K)$, and let $\bar{x} \in \mathcal{K}$ satisfy $d_{\mathcal{K}}(x) = |x - \bar{x}|$. By $(ISP)_r$, there exists $x_0 \in \mathcal{K}$ so that $B := \overline{\mathbb{B}}_r(\bar{x}_0)$ satisfies $\bar{x} \in B \subset K$. Suppose $|z| < \frac{r}{2}$. Note that

$$\begin{aligned} &\bar{d}_{\mathcal{K}}(x+z) + \bar{d}_{\mathcal{K}}(x-z) - 2\bar{d}_{\mathcal{K}}(x) \\ &\qquad\qquad \leq \bar{d}_B(x+z) + \bar{d}_B(x-z) - 2\bar{d}_B(x) \\ &\qquad\qquad \leq C|z|^2, \end{aligned}$$

where $C = C(r, R)$ is the semiconcave constant associated with $\bar{d}_{\mathcal{K}}(\cdot)$ for each convex subset of $\mathbb{B}_R(0) \cap \left[\mathbb{R}^n \setminus \{\mathbb{B}_{\frac{r}{2}}(x_0)\}\right]$. The semiconcave inequality is valid in the last estimate due to the fact that $\|x \pm z - \bar{x}_0\| \geq \frac{r}{2}$.

On the other hand, whenever $|z| \geq \frac{r}{2}$, we have that

$$\bar{d}_{\mathcal{K}}(x+z) + \bar{d}_{\mathcal{K}}(x-z) - 2\bar{d}_{\mathcal{K}}(x) \leq 4R\left(\frac{2r}{2r}\right)^2 \leq \frac{16R}{r^2}|z|^2$$

since $\bar{d}_K(x \pm z) \leq 2R$, and the proof of (8) is thus complete.
In order to prove (9), first we note that (see figure 1)

$$\bar{d}_{\mathcal{K}}(y) \leq \bar{d}_B(y) - \bar{d}_B(x) \tag{10}$$

Since $x \mapsto \bar{d}_B(x)$ is a smooth function (in particular $C^{1,1}$) away from the point x_B, we obtain that

$$\bar{d}_B(y) - \bar{d}_B(x) \leq \left\langle \nabla \bar{d}_B(x), y - x \right\rangle + C|y-x|^2$$

which in turns proves (9). □

It is easy to see that the above result implies Proposition 7. In particular if we consider the case in which $x + z$ and $x - z$ belong in $\mathbb{R}^n \setminus \mathrm{int}\,\mathcal{K}$ then (8) holds with the signed distance function replaced by d_K which in turns implies the semiconcavity. The Proposition 8 allows us to consider cases in which points along a line segment $[x - z, x + z]$ belong to K.

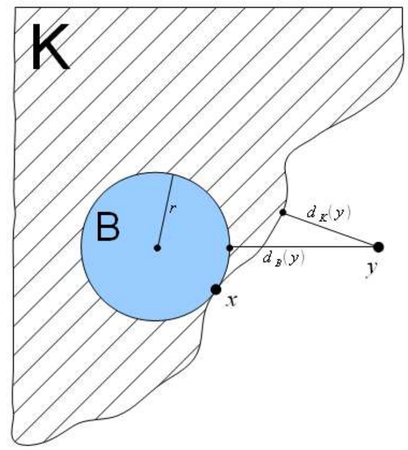

Figure 1: Geometrical meaning of inequality (10)

Remark 2. Another consequence of $(ISP)_r$ is that the transversality condition on the adjoint arc $p(\cdot)$ can be further specified. If $\bar{x}(\cdot)$ is optimal, then by the $(ISP)_r$ assumption, there exists a ball $B := \mathbb{B}_r(x_B)$ contained in $\mathcal{K}$ that touches the boundary of $\mathcal{K}$ at the point $\bar{x}\big(T\big)$ (where $T = T(x)$), and hence $\bar{x}(\cdot)$ is also optimal if the target $\mathcal{K}$ is replaced by B. Applying (NC) to this new target says that we can find an adjoint arc with $p(T) = -\nabla \bar{d}_B\big(\bar{x}(T)\big)$.

3 The main result

We prove in this section that, under suitable assumptions, the minimum time function is locally semiconcave. First we prove the semiconcavity result in a neighborhood of the target set and then we will obtain the local semiconcavity in the whole controllable set.

Theorem 2. *Suppose F satisfies (SH), (H), and that $\mathcal{K}$ is compact. Assume further that the Petrov condition (PC) and the Interior Sphere Property (ISP)$_r$ both hold. Then there exists $\rho > 0$ so that $T(\cdot)$ is semiconcave on each convex set $K \subseteq \mathcal{K}_\rho \setminus \mathcal{K}$ with a constant independent of K.*

Proof. The constant $\rho > 0$ is chosen as in Proposition 6. It suffices to show that there exists a constant $\bar{c} > 0$ so that for every x and $\varepsilon > 0$ with $x + \varepsilon\mathbb{B} \in$

$\mathcal{K}_\rho \setminus \mathcal{K}$, we have

$$T(x+z) + T(x-z) - 2T(x) \leq \bar{c}|z|^2 \tag{11}$$

for all $\|z\| < \varepsilon$. So fix x and $\varepsilon > 0$ so that $x + \varepsilon\mathbb{B} \subseteq \mathcal{K}_\rho \setminus \mathcal{K}$, and fix z with $\|z\| < \varepsilon$. We choose the constant C so that the conclusion of Proposition 8 holds for all vectors that we analyze.

There exists an optimal trajectory $\bar{x}(\cdot)$ starting from x and terminating at $x(T) \in \mathcal{K}$, where $T = T(x)$. By (NC) and Proposition 5, there exists an adjoint arc $p(\cdot)$ (normalized so that $\|p(T)\| = 1$) for which the following differential equation holds.

$$\begin{cases} \dot{\bar{x}}(t) = F_{p(t)}(\bar{x}(t)) & \text{a.e. } t \in [0,T] \\ x(0) = x. \end{cases} \tag{12}$$

With this same $p(\cdot)$, let $x_\pm(\cdot)$ be the solutions of (IVP) with initial conditions $x_\pm(0) = x \pm z$. That is,

$$\begin{cases} \dot{x}_\pm(t) = F_{p(t)}(x_\pm(t)) & \text{a.e. } t \in [0,T] \\ x(0) = x \pm z \end{cases} \tag{13}$$

Recall Proposition 4(2), where a consequence of the above ODEs is that

$$H\big(x(t), p(t)\big) = \big\langle p(t), \dot{x}(t)\big\rangle \quad \text{a.e. } t \in [0,T], \tag{14}$$

where $x(\cdot)$ is any one among the arcs $\bar{x}(\cdot)$, $x_-(\cdot)$, and $x_+(\cdot)$. Another feature of the ODEs is dependence on initial data, and Proposition 4(3) says that

$$|\bar{x}(t) - x_\pm(t)| \leq k\,|z| \quad \text{and} \quad |x_+(t) - x_-(t)| \leq 2\,k\,|z| \quad \forall t \in [0.T]. \tag{15}$$

As a first case, let us assume $x_\pm(t) \notin \mathcal{K}$ for $t \in [0,T]$. To be consistent with notation of the second case below, it is convenient to denote the time when one of the arcs has hit the target by t^*, and so in our case here, we have $t^* = T = T(x)$. We shall frequently have occasion to refer to arc representing teh midpoint between the arcs $x_\pm(\cdot)$, and so is given by

$$\tilde{x}(t) := \frac{1}{2}\big[x_+(t) + x_-(t)\big] \quad \text{for } t \in [0,t^*]. \tag{16}$$

Two different subcases emerge depending on whether $\tilde{x}(t^*) \in \mathcal{K}$ or not.

In the first subcase where $\tilde{x}(t^*) \in \mathcal{K}$, there exists $\bar{t} < t^*$ such that $\tilde{x}(\bar{t}) \in$ bdry $\mathcal{K}$. By the dynamic programming principle, we have

$$\begin{aligned} T(x+z) + T(x-z) - 2T(x) &\leq 2\bar{t} + T\big(x_+(\bar{t})\big) + T\big(x_-(\bar{t})\big) - 2t^* \\ &\leq T\big(x_+(\bar{t})\big) + T\big(x_-(\bar{t})\big) \end{aligned} \tag{17}$$

By Proposition 6, Proposition 8, and (15), we obtain

$$\begin{aligned} T\big(x_+(\bar t)\big) + T\big(x_-(\bar t)\big) &\le m\left[d_{\mathcal K}\big(x_+(\bar t)\big) + d_{\mathcal K}\big(x_-(\bar t)\big)\right] \\ &= m\left[d_{\mathcal K}\left(\tilde x(\bar t) + \frac{x_+(\bar t) - x_-(\bar t)}{2}\right)\right. \\ &\qquad \left. + d_{\mathcal K}\big(\tilde x(\bar t) - \left(\frac{x_+(\bar t) - x_-(\bar t))}{2}\right) - 2d_{\mathcal K}(\tilde x(\bar t))\right] \\ &\le \frac{1}{4}\, m\, C\left|x_+(\bar t) - x_-(\bar t)\right|^2 \\ &\le \bar c |z|^2 \end{aligned}$$

with $\bar c = m\, C\, k^2$. Combining this estimate with (17) yields (11).

The second subcase is where $\tilde x(t^*) \notin \mathcal K$, and by the dynamic programming principle, we have

$$\begin{array}{rcl} T(x) &=& t^*; \\ T(x+z) &\le& t^* + T(x_+(t^*));\ \text{and} \\ T(x-z) &\le& t^* + T(x_-(t^*)). \end{array}$$

Combining these gives

$$T(x+z) + T(x-z) - 2T(x) \le T\big(x_+(t^*)\big) + T\big(x_-(t^*)\big). \tag{18}$$

Now since $d_{\mathcal K}(\cdot)$ bounds $T(\cdot)$ (Proposition 6), we have

$$T(x_+(t^*)) + T(x_-(t^*)) \le m\Big[d_{\mathcal K}\big(x_+(t^*)\big) + d_{\mathcal K}\big(x_-(t^*)\big)\Big], \tag{19}$$

and from Proposition 8 that

$$\begin{aligned} d_{\mathcal K}(x_+(t^*)) + d_{\mathcal K}(x_-(t^*)) &= d_{\mathcal K}(x_+(t^*)) + d_{\mathcal K}(x_-(t^*)) \\ &\qquad -2d_{\mathcal K}(\tilde x(t^*)) + 2d_{\mathcal K}(\tilde x(t^*)) \\ &\le C\,|x_+(t^*) - x_-(t^*)|^2 + 2d_{\mathcal K}(\tilde x(t^*)) \\ &\le 4\,k^2\,C\,|z|^2 + 2d_{\mathcal K}(\tilde x(t^*)) \end{aligned} \tag{20}$$

where the last inequality again follows from (15). We need to obtain the appropriate estimate on $d_{\mathcal K}(\tilde x(t^*))$.

Since $\tilde x(t^*) \notin \mathcal K$, the signed distance function $\bar d_{\mathcal K}\big(\tilde x(t^*)\big)$ is equal to $d_{\mathcal K}\big(\tilde x(t^*)\big)$, and applying remark 2 and proposition 8(2) with $y = \tilde x(t^*)$ and $x = \bar x(t^*)$, we obtain

$$\begin{aligned} d_{\mathcal K}(\tilde x(t^*)) &\le -\big\langle p(t^*), \tilde x(t^*) - \bar x(t^*)\big\rangle + C\,|\tilde x(t^*) - \bar x(t^*)|^2 \\ &\le -\big\langle p(t^*), \tilde x(t^*) - \bar x(t^*)\big\rangle + C\,k^2\,|z|^2 \end{aligned} \tag{21}$$

where the second inequality follows from (15).

The following lemma will assist in getting the appropriate estimate on the inner product term in (21).

Lemma 1. *Let* $\bar{x}(\cdot)$ *and* $p(\cdot)$ *be as above. Suppose* $0 \le t_1 < t_2 \le T(x)$ *and* $x(\cdot)$ *is a solution to (4) with initial condition* $x(t_1)$*. Then*

$$\big\langle p(t_2), \bar{x}(t_2) - x(t_2)\big\rangle - \big\langle p(t_1), \bar{x}(t_1) - x(t_1)\big\rangle \le c\,k^2\,(t_2 - t_1)\,\big|\bar{x}(t_1) - x(t_1)\big|^2.$$

Proof. We have

$$\begin{aligned}\big\langle p(t_2), \bar{x}(t_2) - x(t_2)\big\rangle &- \big\langle p(t_1), \bar{x}(t_1) - x(t_1)\big\rangle \\ &= \int_{t_1}^{t_2} \frac{d}{dt}\big\langle p(t), \bar{x}(t) - x(t)\big\rangle\,dt \\ &= \int_{t_1}^{t_2} \Big[\big\langle \dot{p}(t), \bar{x}(t) - x(t)\big\rangle + \big\langle p(t), \dot{\bar{x}}(t) - \dot{x}(t)\big\rangle\Big]\,dt \qquad (22)\end{aligned}$$

The second integrand in (22) satisfies

$$\big\langle p(t), \dot{\bar{x}}(t) - \dot{x}(t)\big\rangle = H\big(\bar{x}(t), p(t)\big) - H\big(x(t), p(t)\big)$$

by (14). The necessary conditions (NC) have that $-p(t) \in \partial_x H\big(\bar{x}(t), p(t)\big)$, and since $x \mapsto H(x,p)$ is semiconvex (which is (H1)), we have by Proposition 1(4) that the previous quantity is

$$\le \big\langle -p(t), \bar{x}(t) - x(t)\big\rangle + c\,|\bar{x}(t) - x(t)|^2. \qquad (23)$$

Now integrate (23) from t_1 to t_2 and insert this estimate into (22). The first integrand in (22) will cancel with the inner product term in (23), and we are left with

$$\begin{aligned}\big\langle p(t_2), \bar{x}(t_2) - x(t_2)\big\rangle - \big\langle p(t_1), \bar{x}(t_1) - x(t_1)\big\rangle &\le c\int_{t_1}^{t_2}\big|\bar{x}(t) - x(t)\big|^2\,dt \\ &\le c\,k^2\,(t_2 - t_1)\big|\bar{x}(t_1) - x(t_1)\big|^2,\end{aligned}$$

where the last inequality holds by Proposition 4(3). □

Returning now to the inner product in (21), we have

$$\begin{aligned}-\big\langle p(t^*), \tilde{x}(t^*) - \bar{x}(t^*)\big\rangle &= \big\langle p(t^*), \bar{x}(t^*) - \tilde{x}(t^*)\big\rangle \\ &= \frac{1}{2}\Big[\big\langle p(t^*), \bar{x}(t^*) - x_+(t^*)\big\rangle + \big\langle p(t^*), \bar{x}(t^*) - x_-(t^*)\big\rangle\Big]\end{aligned}$$

Recall (13), in which it is stated that both $x_\pm(\cdot)$ satisfy the conditions of Lemma 1. We apply Lemma 1 twice with $t_1 = 0$ and $t_2 = t^*$, and noting

$$\big\langle p(0), \bar{x}(0) - x_+(0)\big\rangle + \big\langle p(0), \bar{x}(0) - x_-(0)\big\rangle = \big\langle p(0), z - z\big\rangle = 0,$$

deduce that

$$\big\langle p(t^*), \bar{x}(t^*) - \tilde{x}(t^*)\big\rangle \le c\,k^2\,t^*\,\big|z\big|^2 \qquad (24)$$

Finally, combining the estimates in (21) and (24) we arrive at

$$d_{\mathcal{K}}\big(\tilde{x}(t^*)\big) \leq k^2\,(C + c\,t^*)\;|z|^2.$$

In conjunction with (20), (18), and (19), we may conclude that (11) holds for $\bar{c} = 2m\,k^2\,(3\,C + c\,t^*)$. This finishes the proof of the first case where $x_\pm(t) \notin \mathcal{K}$ for all $t \in [0, T]$

We now turn to the case where one of $x_\pm(t)$ hits $\mathcal{K}$ before $\bar{x}(\cdot)$ does. For definiteness, suppose $x_+(\cdot)$ reaches the target before the others, and so $x_+(t^*) \in \mathcal{K}$ for some $t^* < T(x)$ and $\bar{x}(t)$ and $x_-(t)$ do not belong to $\mathcal{K}$ for all $t \in [0, t^*]$. The other case is similar where $x_-(\cdot)$ hits first is similar. The case we are currently analyzing is portrayed in Figure 2.

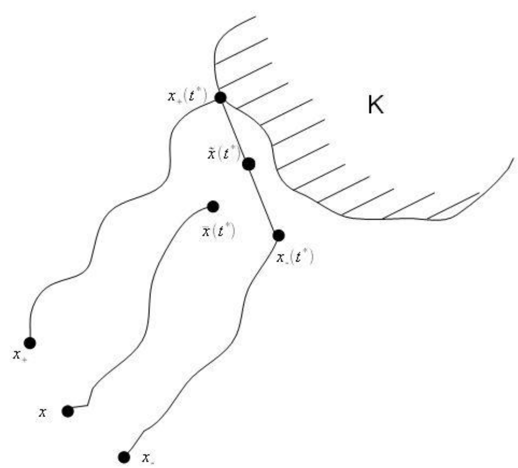

Figure 2: Trajectories up to time t^*

The dynamic programming principle implies

$$\begin{aligned} T(x+z) &\leq t^*, \\ T(x-z) &\leq t^* + T\big(x_-(t^*)\big), \text{ and} \\ T(x) &= t^* + T\big(\bar{x}(t^*)\big). \end{aligned}$$

Therefore with $\tilde{x}(\cdot)$ defined as in (16), we have

$$\begin{aligned} &T(x+z) + T(x-z) - 2T(x) \\ &\qquad \leq T\big(x_-(t^*)\big) - 2T\big(\bar{x}(t^*)\big) \\ &\qquad = \Big[T\big(x_-(t^*)\big) - 2T\big(\tilde{x}(t^*)\big)\Big] + 2\Big[T\big(\tilde{x}(t^*)\big) - T\big(\bar{x}(t^*)\big)\Big] \end{aligned} \tag{25}$$

To prove (11), we need to show both of the bracketed terms in (25) are bounded above by $\bar{c}\,|z|^2$. The proof of each estimate is somewhat lengthy.

We consider the second term first. That is, we will show

$$T\big(\tilde{x}(t^*)\big) - T\big(\bar{x}(t^*)\big) \le \bar{c}\,|z|^2. \tag{26}$$

Let $\hat{x}(\cdot)$ be the solution of

$$\begin{cases} \dot{\hat{x}}(t) = F_{p(t)}\big(\hat{x}(t)\big) & \text{a.e. } t \in [t^*, T(x)] \\ \hat{x}(t^*) = \tilde{x}(t^*), \end{cases}$$

which (unlike $\tilde{x}(\cdot)$) is a trajectory of (2). If $\hat{x}(\cdot)$ reaches $\mathcal{K}$ before $\bar{x}(\cdot)$ does, then $T(\tilde{x}(t^*)) < T(\bar{x}(t^*))$ and (26) is trivial. So suppose $\hat{x}(t) \notin \mathcal{K}$ for $t \in [t^*, T(x)]$, and set $\bar{t} := T(x) - t^*$. By the dynamic programming principle, we have

$$\begin{aligned} &T(\bar{x}(t^*)) = \bar{t} \\ &T\big(\tilde{x}(t^*)\big) \le \bar{t} + T\big(\hat{x}(\bar{t})\big), \end{aligned}$$

and therefore

$$T\big(\tilde{x}(t^*)\big) - T\big(\bar{x}(t^*)\big) \le T\big(\hat{x}(\bar{t})\big) \le m d_{\mathcal{K}}\big(\hat{x}(\bar{t})\big) \tag{27}$$

by Proposition 6.
Applying remark 2 and proposition 8(2) with $y = \hat{x}(\bar{t})$ and $x = \bar{x}(\bar{t})$, the following estimate holds

$$d_{\mathcal{K}}\big(\hat{x}(\bar{t})\big) \le -\big\langle p(\bar{t}), \hat{x}(\bar{t}) - \bar{x}(\bar{t})\big\rangle + C\big|\hat{x}(\bar{t}) - \bar{x}(\bar{t})\big|^2. \tag{28}$$

Similar to the inequality in (15), we have by Proposition 4(3), and then using (15) twice, that

$$\big|\hat{x}(t) - \bar{x}(t)\big| \le k\big|\hat{x}(t^*) - \bar{x}(t^*)\big| = \frac{k}{2}\big|x_+(t^*) + x_-(t^*) - 2\bar{x}(t^*)\big| \le k^2|z| \tag{29}$$

for all $t \in [t^*, T(x)]$. Substituting (28)and (29) into (27) yields

$$T(\tilde{x}(t^*)) - T(\bar{x}(t^*)) \le m\,\big\langle p(\bar{t}), \bar{x}(\bar{t}) - \hat{x}(\bar{t})\big\rangle + m\,k^4\,C\,|z|^2 \tag{30}$$

To estimate the first term in the right hand side of (30), we invoke Lemma 1 with $t_1 = t^*$ and $t_2 = \bar{t}$, and deduce

$$\begin{aligned} &\big\langle p(\bar{t}), \bar{x}(\bar{t}) - \hat{x}(\bar{t})\big\rangle \\ &\qquad \le \big\langle p(t^*), \bar{x}(t^*) - \hat{x}(t^*)\big\rangle + c\,k^2\,\big|\bar{x}(t^*) - \hat{x}(t^*)\big|^2 \\ &\qquad = \frac{1}{2}\Big[\big\langle p(t^*), \bar{x}(t^*) - x_+(t^*)\big\rangle + \big\langle p(t^*), \bar{x}(t^*) - x_-(t^*)\big\rangle\Big] \\ &\qquad\qquad + c\,k^2\,\big|\bar{x}(t^*) - \tilde{x}(t^*)\big|^2 \end{aligned}$$

The two terms in the brackets can also be estimated using Lemma 1 with $t_1 = 0$ and $t_2 = t^*$. The initial values make no contribution this time since $\bar{x}(0) = \frac{1}{2}\big(x_+(0) + x_-(0)\big)$. Therefore the bracketed term is bounded by $c\,k^2\,t^*\,|z|^2$, and the last term is bounded by $c\,k^4\,|z|^2$. Substituting all of this into (30) yields (26) with $\bar{c} := m\,k^2\,\big(c\,t^* + c\,k^2 + C\,k^2\big)$.

The last consideration to finish the proof of Theorem 2 is to appropriately bound the first term in (25). We must show

$$T(x_-(t^*)) - 2T(\tilde{x}(t^*)) \leq \bar{c}\,|z|^2. \tag{31}$$

It is convenient at this point to introduce new notation because the estimates do not directly involve x and z, but rather depend on the values of the trajectories $x_\pm(\cdot)$ at time t^*. We define y and w by setting

$$y := x_+(t^*) \quad \text{and} \quad w := \tilde{x}(t^*) - x_+(t^*) = \frac{x_-(t^*) - x_+(t^*)}{2},$$

and note that $y \in \mathcal{K}$, $y + w = \tilde{x}(t^*)$, and $y + 2\,w = x_-(t^*)$. Since $|w| \leq k|z|$ by (15), in order to prove (31), it suffices to show

$$T(y + 2w) - 2T(y + w) \leq \bar{c}\,|w|^2. \tag{32}$$

Toward this end, let $\bar{y}(\cdot)$ be optimal solution to the minimal time problem starting from $y + w$, and now denote by $p(\cdot)$ its associated adjoint arc. If $\theta^\star := T(y + w)$ satisfies $T(y + 2w) \leq 2\theta^\star$, then (32) is trivial, so suppose this is not the case, and let $y(\cdot)$ be the solution of

$$\begin{cases} \dot{y}(t) = F_{p(\frac{t}{2})}(y(t)) & \text{a.e. } t \in [0, 2\theta^\star] \\ y(0) = y + 2w. \end{cases}$$

The dynamic programming principle says

$$T(y + 2w) \leq 2\theta^* + T\big(y(2\theta^\star)\big),$$

and therefore (32) follows upon showing that $T\big(y(2\theta^\star)\big)$ is bounded by $\bar{c}|w|^2$. We have by Proposition 6 that

$$\begin{aligned} T\big(y(2\theta^*)\big) &\leq m\,d_{\mathcal{K}}\big(y(2\theta^*)\big) \\ &\leq m\Big[\big|y(2\theta^*) - 2\,\bar{y}(\theta^*) + y\big| + d_{\mathcal{K}}\big(2\bar{y}(\theta^*) - y\big)\Big] \\ &= m\Big[(I)\; + \;(II)\Big]. \end{aligned} \tag{33}$$

To bound (I), note the fundamental theorem of calculus, a change of variables, and assumption (H2) imply

$$
\begin{aligned}
(I) &= \left| \int_0^{2\theta^*} \dot{y}(t)dt - 2\int_0^{\theta^*} \dot{\bar{y}}(t)dt \right| \\
&= 2\left| \int_0^{\theta^*} \left[F_{p(t)}(y(2t)) - F_{p(t)}(\bar{y}(t))\right] dt \right| \\
&\leq 2\ell \int_0^{\theta^*} |y(2t) - \bar{y}(t)| dt
\end{aligned}
\tag{34}
$$

where ℓ is the Lipschitz constant contained in the assumption (H2). By standard ODE theory, there exists a constant k so that for all $0 \leq t_1,\, t_2 \leq \theta^*$ so that

$$
|y(2\,t_1) - \bar{y}(t_2)| \leq k\,|w|. \tag{35}
$$

We also record the fact that

$$
\theta^* \leq m\,|w|, \tag{36}
$$

which is true because $\theta^* = T(y+w) \leq m\, d_{\mathcal{K}}(y+w) \leq m\,|w|$ since $y \in \mathcal{K}$. Consequently, using (35) and (36) to estimate (34), we conclude the term (I) is bounded by $2\,\ell\, m\, k\,|w|^2$.

With regard to the second term of the (33), recall that y and $\bar{y}(\theta^*)$ belong to $\mathcal{K}$, and then note the semiconcavity hypothesis of the distance function implies

$$
\begin{aligned}
d_{\mathcal{K}}\big(2\bar{y}(\theta^*) - y\big) &= d_{\mathcal{K}}\big(2\bar{y}(\theta^*) - y\big) + d_{\mathcal{K}}(y) - 2\, d_{\mathcal{K}}\big(\bar{y}(\theta^*)\big) \\
&\leq C\, \big|\bar{y}(\theta^*) - y\big|^2.
\end{aligned}
\tag{37}
$$

We have

$$
|\bar{y}(\theta^*) - y| \leq \left| w + \int_0^{\theta^*} \dot{\bar{y}}(t)\, dt \right| \leq |w| + k\theta^* \leq (1+km)|w|
$$

by (36). Hence (37) now implies (II) is bounded above by $(1+km)^2\, C\,|w|^2$.

Finally, combining all of the estimates, we see that (32) holds for $\bar{c} = m\,\big(2\,\ell\, m\, k + (1+km)^2\, C\,\big)$.

This completes the proof of the last case, and indeed of the the whole theorem. □

Roughly speaking the above theorem tells us that the minimum time function is semiconcave in a neighborhood of the target set. This is a preliminary result useful for proving the local semiconcavity in the whole controllable set $\mathcal{R}\setminus\mathcal{K}$. Since the most difficult work was in the Theorem 2, the local semiconcavity result is given as a consequence. Indeed by the dynamic programming principle we obtain the following result.

Corollary 1. *Under the same assumptions of the Theorem 2, the minimum time function is locally semiconcave on $\mathcal{R}\setminus\mathcal{K}$.*

Proof. Let C be any compact set of $\mathcal{R}\setminus\mathcal{K}$. Fix $x\in C$ and $\varepsilon>0$ such that $x\pm\varepsilon\mathbb{B}\in C$. Moreover we fix z with $\|z\|<\varepsilon$ and such that there exists t^* in which the trajectories defined as in the Theorem 2 (13) and (12) belong to the set $\mathcal{K}_\rho\setminus\mathcal{K}$ with $\rho>0$ fixed as in Theorem 2.
By the dynamic programming principle

$$\begin{aligned}&T(x+z)+T(x-z)-2T(x)\leq\\&\leq T(x_+(t^*))+T(x_-(t^*))-2T(\tilde{x}(t^*))+2\big[T(\tilde{x}(t^*))-T(\bar{x}(t^*))\big]\end{aligned}\tag{38}$$

where $\tilde{x}(t^*)$ is defined as (16). Applying the semiconcavity result on a neighborhood of the target set we can find a constant $c>0$ such that

$$T(x+z)+T(x-z)-2T(x)\leq c|z|^2+2\big[T(\tilde{x}(t^*))-T(\bar{x}(t^*))\big]$$

As in (26) let $\hat{x}(\cdot)$ be the solution of

$$\begin{cases}\dot{\hat{x}}(t)=F_{p(t)}\big(\hat{x}(t)\big) & \text{a.e. } t\in[t^*,T(x)]\\ \hat{x}(t^*)=\tilde{x}(t^*),\end{cases}$$

which is a trajectory of (2). We can consider the only case in which $\hat{x}(t)\notin\mathcal{K}$ for $t\in[t^*,T(x)]$ (otherwise it is trivial), and set $\bar{t}:=T(x)-t^*$. By the dynamic programming principle, we have

$$\begin{aligned}&T(\bar{x}(t^*))=\bar{t}\\&T\big(\tilde{x}(t^*)\big)\leq\bar{t}+T\big(\hat{x}(\bar{t})\big),\end{aligned}$$

and therefore, by Proposition 6

$$T\big(\tilde{x}(t^*)\big)-T\big(\bar{x}(t^*)\big)\leq T\big(\hat{x}(\bar{t})\big)\leq md_{\mathcal{K}}\big(\hat{x}(\bar{t})\big)\tag{39}$$

Applying remark 2 and proposition 8(2) with $y=\hat{x}(\bar{t})$ and $x=\bar{x}(\bar{t})$, the following estimate holds

$$d_{\mathcal{K}}\big(\hat{x}(\bar{t})\big)\leq-\big\langle p(\bar{t}),\hat{x}(\bar{t})-\bar{x}(\bar{t})\big\rangle+C\big|\hat{x}(\bar{t})-\bar{x}(\bar{t})\big|^2.\tag{40}$$

using (15) twice and thanks to Proposition 4(3) we have

$$\big|\hat{x}(t)-\bar{x}(t)\big|\leq k\big|\hat{x}(t^*)-\bar{x}(t^*)\big|=\frac{k}{2}\big|x_+(t^*)+x_-(t^*)-2\bar{x}(t^*)\big|\leq k^2|z|\tag{41}$$

for all $t\in[t^*,T(x)]$. Substituting (40)and (41) into (39) yields

$$T(\tilde{x}(t^*))-T(\bar{x}(t^*))\leq m\,\big\langle p(\bar{t}),\bar{x}(\bar{t})-\hat{x}(\bar{t})\big\rangle+m\,k^4C\,|z|^2\tag{42}$$

Lemma 1 with $t_1 = t^*$ and $t_2 = \bar{t}$ implies that

$$\begin{aligned}\big\langle p(\bar{t}), \bar{x}(\bar{t}) - \hat{x}(\bar{t})\big\rangle \\ &\leq \big\langle p(t^*), \bar{x}(t^*) - \hat{x}(t^*)\big\rangle + c\,k^2 \left|\bar{x}(t^*) - \hat{x}(t^*)\right|^2 \\ &= \frac{1}{2}\Big[\big\langle p(t^*), \bar{x}(t^*) - x_+(t^*)\big\rangle + \big\langle p(t^*), \bar{x}(t^*) - x_-(t^*)\big\rangle\Big] \\ &\qquad + c\,k^2 \left|\bar{x}(t^*) - \tilde{x}(t^*)\right|^2\end{aligned}$$

Using Lemma 1 again with $t_1 = 0$ and $t_2 = t^*$ we can find a constant $\bar{c} := m\,k^2\left(c\,t^* + c\,k^2 + k^2 C\right)$ such that

$$T(\tilde{x}(t^*)) - T(\bar{x}(t^*)) \leq \bar{c}|z|^2$$

which in turn implies the semiconcavity property on the whole controllable set $\mathbb{R} \setminus \mathcal{K}$. □

References

[1] J-P. Aubin and A. Cellina. *Differential inclusions.* Springer-Verlag, Berlin, 1984. 6

[2] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia. 2

[3] M. Bardi and M. Falcone. An approximation scheme for the minimum time function. *SIAM J. Control Optim.*, 28(4):950–965, 1990. 2

[4] P. Cannarsa and H. Frankowska. Some characterizations of optimal trajectories in control theory. *SIAM J. Control Optim.*, 29(6):1322–1347, 1991. 3

[5] P. Cannarsa and H. Frankowska. Interior sphere property of attainable sets and time optimal control problems. *ESAIM Control Optim. Calc. Var.*, 12(2):350–370 (electronic), 2006. 2

[6] P. Cannarsa, H. Frankowska, and C. Sinestrari. Optimality conditions and synthesis for the minimum time problem. *Set-Valued Anal.*, 8(1-2):127–148, 2000. Set-valued analysis in control theory. 2, 3

[7] P. Cannarsa and C. Sinestrari. Convexity properties of the minimum time function. *Calc. Var. Partial Differential Equations*, 3(3):273–298, 1995. 2, 8

[8] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston Inc., Boston, MA, 2004. 2, 3, 5, 8

[9] P. Cannarsa and P.R. Wolenski. Semiconcavity of the value function for a class of differential inclusions. *Discrete Contin. Dyn. Syst. Ser. B*, 29(2):453–466, 2011. 2, 3, 6, 7

[10] F.H. Clarke. *Optimization and nonsmooth analysis*. Wiley-Interscience, New York, 1983. 2, 6

[11] F.H. Clarke, Yu. Ledyaev, R.J. Stern, and P.R. Wolenski. *Nonsmooth analysis and control theory*. Springer-Verlag, New York, 1998. 2, 3, 6

[12] H. Hermes and J. P. LaSalle. *Functional analysis and time optimal control*. Academic Press, New York, 1969. Mathematics in Science and Engineering, Vol. 56. 2

[13] N. N. Petrov. The Bellman problem for a time-optimality problem. *Prikl. Mat. Meh.*, 34:820–826, 1970. 2

[14] N.N. Petrov. On the bellman function for the time-optimal process problem. *Journal of Applied Mathematics and Mechanics*, 34(5):785–791, 1970. 2

[15] R. T. Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998. 5

[16] C. Sinestrari. Semiconcavity of the value function for exit time problems with nonsmooth target. *Commun. Pure Appl. Anal.*, 3(4):757–774, 2004. 2

[17] P. Soravia. Hölder continuity of the minimum-time function for C^1-manifold targets. *J. Optim. Theory Appl.*, 75(2):401–421, 1992. 2

[18] V. M. Veliov. Lipschitz continuity of the value function in optimal control. *J. Optim. Theory Appl.*, 94(2):335–363, 1997. 2

[19] P.R. Wolenski and Y. Zhang. Proximal analysis and the minimal time function. *SIAM J. Control Optim.*, 36(3):1048–1072, 1998. 2

Received ; revised .

http://monotone.uwaterloo.ca/~journal/