HAL open science

Conjugate times and regularity of the minimum time function with differential inclusions

Piermarco Cannarsa, Teresa Scarinci

To cite this version:

Piermarco Cannarsa, Teresa Scarinci. Conjugate times and regularity of the minimum time function with differential inclusions. 2015. hal-01100971v2

HAL Id: hal-01100971

https://hal.sorbonne-universite.fr/hal-01100971v2

Preprint submitted on 21 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conjugate times and regularity of the minimum time function with differential inclusions

Piermarco Cannarsa and Teresa Scarinci

Abstract This paper studies the regularity of the minimum time function, $T(\cdot)$, for a control system with a closed target, taking the state equation in the form of a differential inclusion. Our first result is a sensitivity relation which guarantees the propagation of the proximal subdifferential of T along any optimal trajectory. Then, we obtain the local C^2 regularity of the minimum time function along optimal trajectories by using such a relation to exclude the presence of conjugate times.

Introduction

This paper aims to refine the study of the regularity properties of the value function of the time optimal control problem in nonparameterized form, that is, when the state equation is given as a differential inclusion. This problem seems hard to address by parametrization techniques, as it has been observed in the recent papers [10], [12], [11], and [9].

Recall the minimum time problem $\mathscr{P}(x)$ 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 $\mathscr{K} \subseteq \mathbb{R}^n$. Specifically, the problem $\mathscr{P}(x)$ is

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

Piermarco Cannarsa
Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy, e-mail: cannarsa@mat.uniroma2.it

Teresa Scarinci
Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy and CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, Univ Paris Diderot, Sorbonne Paris Cité, Case 247, 4 Place Jussieu, 75252 Paris, France, e-mail: teresa.scarinci@gmail.com

where the minimization is over all absolutely continuous arcs $y(\cdot)$ defined on an interval $[0,T]$ that satisfy the differential inclusion

$$\begin{cases} \dot{y}(t) \in F\big(y(t)\big) & \text{a.e. } t \in [0,T] \\ y(0) = x \end{cases} \tag{2}$$

and the terminal condition $y(T) \in \mathscr{K}$. Here, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a Lipschitz continuous multifunction having a sublinear growth such that the associated Hamiltonian

$$H(x,p) = \sup_{v \in F(x)} \langle -v, p \rangle \qquad (x,p) \in \mathbb{R}^n \times \mathbb{R}^n$$

is semiconvex in x and differentiable in p, whenever $p \neq 0$. The minimum time function, $T(x)$, is defined as the optimal value in (1).

The main object of our analysis are sensitivity relations, that is, inclusions that identify the dual arc as a suitable generalized gradient of the minimum time function $T(\cdot)$, evaluated along a given minimizing trajectory. The importance of such relations is well acknowledged and will be made clear by the applications we provide to the differentiability of $T(\cdot)$.

Sensitivity relations have a long history dating back, at least, to the papers [17], [23], [2], and [22] that studied optimal control problems of Bolza type with finite time horizon. In [8], such relations were adapted to the minimum time problem for the parameterized control system

$$\dot{y}(t) = f(y(t), u(t)) \quad t \geq 0 \tag{3}$$

assuming that:

(i) $\mathscr{K}$ has the inner sphere property, and
(ii) Petrov's controllability condition is satisfied on $\partial\mathscr{K}$.

For any optimal trajectory $y(\cdot)$ of (3) originating at a point x in the controllable set, the result of [8] ensures the existence of an arc p, called a dual arc, such that:

- (y,p) satisfies the Hamiltonian system

$$\begin{cases} -\dot{y}(t) = \nabla_p H(y(t), p(t)) \\ \dot{p}(t) \in \partial_x^- H(y(t), p(t)) \end{cases} \quad 0 \leq t \leq T(x) =: T \tag{4}$$

  together with the *transversality* condition

$$p(T) = -\frac{\nu}{H(y(T), -\nu)},$$

  where ν is any unit *inner* normal to $\mathscr{K}$ at $y(T)$;
- $p(t)$ belongs to the *Fréchet* superdifferential of $T(\cdot)$ at $y(t)$ for all $t \in [0, T(x))$.

In [11], the above result was extended to nonparameterized control systems by developing an entirely different proof, based on the Pontryagin maximum principle

rather than linearization techniques as in [8]. In [9], assumption (ii) above was removed, still keeping (i) in force, showing that $p(t)$ is either a *proximal* or a *horizontal* supergradient of $T(\cdot)$ at $y(t)$, for all $t \in [0, T(x))$, depending on whether Petrov's condition is satisfied or not at $y(T(x))$.

With respect to sensitivity relations, the purpose of the present paper is to derive analogous inclusions for the the *proximal subdifferential* of the minimum time function; more precisely, we will prove the propagation of the subdifferential of $T(\cdot)$ along optimal trajectories. By 'propagation of the proximal subdifferential' we mean the fact that, if a proximal subgradient of $T(\cdot)$ exists at some point x of the reachable set—so that the minimum time function is differentiable at x—and $y(\cdot)$ is a time optimal trajectory starting at x, then $p(t)$ belongs to the proximal subgradient of $T(\cdot)$ at $y(t)$ for all $t \in [0, T(x))$. Such an invariance of the subdifferential with respect to the Hamiltonian flow associated with (4) was pointed out in [4] for functionals in the calculus of variations and [5] for optimal control problems of Bolza type. A similar result was obtained in [6] for the Mayer problem and in [18] for the minimum time problem for parameterized control system. In Theorem 3 of this paper, we show that such a property holds for the minimum time problem with a state equation in the form of a differential inclusion.

We give two applications of the above sensitivity relations. The first one (Theorem 5 below) ensures that the differentiability of $T(\cdot)$ propagates along an optimal trajectory, $y(\cdot)$, originating at a point x of the controllable set if and only if Petrov's condition is satisfied at $y(T(x))$. This property follows directly from the above relations which guarantee that the corresponding dual arc is contained in both Fréchet semidifferentials whenever $T(\cdot)$ is differentiable at x . Our second application concerns the local smoothness of the minimum time function along an optimal trajectory $y(\cdot)$, that is, the property of having continuous second order derivatives in a neighborhood of $\{y(t) \ : \ 0 \leq t < T(x)\}$. In Theorem 6, we show that this is indeed the case whenever $T(\cdot)$ has a proximal subgradient at the starting point of $y(\cdot)$.

In order to prove the local smoothness of $T(\cdot)$ along an optimal trajectory we need to analyze conjugate times, and give sufficient conditions to exclude the presence of such times. The notion of conjugate point is classical in the calculus of variations and optimal control. Recently, conjugate times have been considered in [18] linearizing the system on the whole $\mathbb{R}^n$ but neglecting the role of the time variable. In such a paper, the degeneracy condition is assigned on the tangent space to the target, which is an $(n-1)$-dimensional space and the authors show that the absence of conjugate times at a point x ensures the C^1-smoothness of $T(\cdot)$ along the trajectory originating at x. In this paper, we return to the 'classical' definition of conjugate point and formulate a sufficient condition for smoothness in terms of conjugate times (see Theorem 2), much in the spirit of the result of [21].

In this way we deduce that, if the proximal subgradient of $T(\cdot)$ is nonempty at some point x, then the minimum time function is locally smooth along the optimal trajectory originating at x.

The paper is organized as follows. Background material is collected in Section 1. In Section 2, we recall preliminary results and discuss the main assumptions we

work with. Section 3 is devoted to the analysis of conjugate times. Section 4 contains our sensitivity relations and their applications to regularity.

1 Notation

Let us fix the notation and list some basic facts. Further details can be found in several books, for instance [1, 16, 24, 13].

We denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^n$ and by $\langle\cdot,\cdot\rangle$ the inner product. $B(x,\varepsilon)$ is the closed ball of radius $\varepsilon > 0$ centered at x, and S^{n-1} the unit sphere in $\mathbb{R}^n$. $\mathbb{R}^{n\times n}$ is the set of $n\times n$ real matrices and $\| Q \|$ is the operator norm of a matrix Q, Q^* is the transpose of Q, $\ker Q$ is the kernel of Q, while I_n is the $n\times n$ identity matrix. Recall that $\| Q \| = \sup\{|\langle Ax,x\rangle| : x \in S^{n-1}\}$ for any symmetric $n\times n$ real matrix Q. Moreover, $\mathrm{co}\,E$, ∂E, $\overline{E}$ and E^C are the convex hull, the boundary, the closure and the complement of a set $E \subset \mathbb{R}^n$, respectively.

Let K be a closed subset of $\mathbb{R}^n$ and $x \in K$. $N_K^C(x)$ denotes the Clarke normal cone to K at x. A vector $v \in \mathbb{R}$ is a *proximal (outer) normal* to K at x, and we write $v \in N_K^P(x)$, if there exists $\sigma = \sigma(x,v)$ such that, for all $y \in K$,

$$\langle v, y-x\rangle \le \sigma |y-x|^2. \tag{5}$$

In K is a convex subset of $\mathbb{R}^n$, the proximal normal cone to K at x coincides with the convex normal cone to K at x.
A vector ν is a proximal inner normal to K at x if $\nu \in N_{K^C}^P(x)$. We call it an unit proximal inner normal to K at x if it belongs to the set $N_{K^C}^P(x) \cap S^{n-1}$.
We say that K satisfies the *inner sphere property* of radius R, $R > 0$, if for every $x \in \partial K$ there exists a nonzero vector $\nu_x \in N_{K^C}^P(x)$ such that (5) holds true with $\sigma = |\nu_x|(2R)^{-1}$ and $v = \nu_x$ and for all $y \in K^C$. Equivalently, for all $x \in \partial K$ there exists a vector $0 \neq \nu_x \in N_{K^C}^P(x)$ *realized by a ball* of radius R, that is,

$$B\left(x + R\frac{\nu_x}{|\nu_x|}, R\right) \subset K.$$

Roughly speaking, if K satisfies the inner sphere property of radius R then we have an upper bound for the curvature of ∂K, even though ∂K may be a nonsmooth set. Indeed, any $x \in \partial K$ belongs to some closed ball $y_x + RB(0,1) \subset K$. This fact suggests that, in some sense, the curvature of ∂K is bounded above and excludes the presence of outward pointing corners on ∂K.

If $f : [t_0,t_1] \to \mathbb{R}^n$ is continuous, $f \in C([t_0,t_1])$, define $\|f\|_\infty = \max_{t\in[t_0,t_1]} |f(t)|$. Moreover, we usually refer to an absolutely continuous function $x : [t_0,t_1] \to \mathbb{R}^n$ as an arc. The space $C^k(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$, is the space of all

functions that are continuously differentiable k times on Ω. The gradient of f is $\nabla f(\cdot)$, if it does exist. Moreover, if f is twice differentiable at some $x \in \Omega$, then $\nabla^2 f(x)$ denotes the Hessian of f at x. Let $f : \Omega \to \mathbb{R}$ be any real-valued function defined on an open set $\Omega \subset \mathbb{R}^n$. Let $x \in \Omega$ and $p \in \mathbb{R}^n$. We say that:

- p is a *Fréchet subgradient* of f at x, $p \in \partial^- f(x)$, if

$$\liminf_{y \to x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|\, y - x \,|} \geq 0,$$

- p is a *Fréchet supergradient* of f at x, $p \in \partial^+ f(x)$, if

$$\limsup_{y \to x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|\, y - x \,|} \leq 0,$$

- p is a *proximal subgradient* of f at x, $p \in \partial^{-,P} f(x)$, if $\exists\, c,\ \rho \geq 0$ such that

$$f(y) - f(x) - \langle p, y - x \rangle \geq -c|y - x|^2,\ \forall y \in B(x, \rho),$$

- p is a *proximal supergradient* of f at x, $p \in \partial^{+,P} f(x)$, if $\exists\, c,\ \rho \geq 0$ such that

$$f(y) - f(x) - \langle p, y - x \rangle \leq c|y - x|^2,\ \forall y \in B(x, \rho).$$

If f is Lipschitz, $\zeta \in \mathbb{R}^n$ is a *reachable gradient* of f at $x \in \Omega$ if there exists a sequence $\{x_j\} \subset \Omega$ converging to x such that f is differentiable at x_j for all $j \in \mathbb{N}$ and $\zeta = \lim_{j \to \infty} \nabla f(x_j)$. Let $\partial^* f(x)$ denote the set of all reachable gradients of f at x. The *(Clarke) generalized gradient* of f at $x \in \Omega$, $\partial f(x)$, is the set $\mathrm{co}\,(\partial^* f(x))$.
For an open set $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ is *semiconcave* if it is continuous in Ω and there exists a constant c such that

$$f(x + h) + f(x - h) - 2f(x) \leq c|h|^2,$$

for all $x, h \in \mathbb{R}^n$ such that $[x - h, x + h] \subset \Omega$. We say that a function f is semiconvex on Ω if and only if $-f$ is semiconcave on Ω. We recall below some properties of semiconcave functions (for further details see, for instance, [13]).

Proposition 1. *Let $\Omega \subset \mathbb{R}^n$ be open, $f : \Omega \to \mathbb{R}$ be a semiconcave function with semiconcavity constant c, and let $x \in \Omega$. Then, f is locally Lipschitz on Ω and the following holds true*

1. *$p \in \mathbb{R}^n$ belongs to $\partial^+ f(x)$ if and only if, for any $y \in \Omega$ such that $[y, x] \subset \Omega$,*

$$f(y) - f(x) - \langle p, y - x \rangle \leq c|y - x|^2. \tag{6}$$

2. *$\partial f(x) = \partial^+ f(x) = co\ (\partial^* f(x))$.*
3. *If $\partial^+ f(x)$ is a singleton, then f is differentiable at x.*

If f is semiconvex, then (6) holds reversing the inequality and the sign of the quadratic term, and the other two statements are true with the subdifferential in-

stead of the superdifferential.

Let $M \subset \mathbb{R}^n$ be a C^m-*manifold* of dimension $n-1$ and fix $\xi_0 \in M$. Let $A \subset \mathbb{R}^{n-1}$ be an open set, let $\phi : A \to \mathbb{R}^n$ be a map of class C^m such that $\phi(A) \subset M$, $D\phi(y)$ has rank equal to $n-1$ for all $y \in A$ and $\phi(\eta_0) = \xi_0$ for some $\eta_0 \in A$. We call ϕ a local *parameterization* of M near ξ_0. The components $(\eta_1, ...\eta_{n-1})$ of a point $\eta = \phi^{-1}(\xi) \in A$ are usually called *local coordinates* of $\xi \in M$.
An application $F : M \to \mathbb{R}^n$ is of class C^k at $\xi_0 \in M$ if the map $F \circ \phi^{-1} : \phi(A) \to \mathbb{R}^n$ is of class C^k at $v_0 := \phi^{-1}(\xi_0)$ for any local parameterization ϕ of M near ξ_0. Equivalently, $F : M \to \mathbb{R}^n$ is of class C^k at ξ_0 if there exists a local parameterization ϕ of M near ξ_0 such that $F \circ \phi^{-1}$ is of class C^k at $\eta_0 := \phi^{-1}(\xi_0)$.

2 Assumptions and preliminary results

The *minimum time problem* $\mathscr{P}(x)$ consists of minimizing the time T over all trajectories of a differential inclusion that start from an initial point $x \in \mathbb{R}^n$ and reach a nonempty compact set $\mathscr{K} \subseteq \mathbb{R}^n$, usually called *target*. Specifically, for any absolutely continuous function $y^x(\cdot) \in AC([0,+\infty);\mathbb{R}^n)$ that solves the differential inclusion

$$\begin{cases} \dot{y}(t) \in F\big(y(t)\big) & \text{a.e. } t \geq 0 \\ y(0) = x, \end{cases} \tag{7}$$

let us denote by

$$\theta(y^x(\cdot)) := \inf\{t \geq 0 : \ y^x(t) \in \mathscr{K}\}$$

the first time at which the trajectory $y^x(\cdot)$ reaches the target $\mathscr{K}$ starting from x. By convention, we set $\theta(y^x(\cdot)) = +\infty$ whenever $y^x(\cdot)$ does not reach $\mathscr{K}$. Here and throughout the paper, $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a given multifunction that satisfies the so-called *Standing Hypotheses*:

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

The *minimum time* function $T : \mathbb{R}^n \to [0,+\infty]$ is defined by: for all $x \in \mathbb{R}^n$,

$$T(x) := \inf\{\theta(y^x(\cdot)) : \ y^x(\cdot) \text{ solves (7)}\}. \tag{8}$$

$T(x)$ represents the minimum time needed to steer the point x to the target $\mathscr{K}$ along the trajectories of (7). It is well-known that (SH) guarantees the existence of absolutely continuous solutions to (7) defined on $[0,+\infty)$. Moreover, if x is in the *reachable set* $\mathscr{R}$ (i.e. $T(x) < +\infty$) then $\mathscr{P}(x)$ has an optimal solution, that is, a solution to (7) that gives the minimum in (8). The main assumptions of this paper are expressed in terms of the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ associated to F, that is, the function defined by

$$H(x,p) = \sup_{v\in F(x)} \langle -v, p\rangle. \tag{9}$$

We shall suppose that

$$(H)\begin{cases} \text{for every } r>0 \\ (i)\ \exists\, c\geq 0 \text{ so that }, \forall p\in S^{n-1}, x\mapsto H(x,p) \text{ is semiconvex on } B(0,r) \text{ with} \\ \quad\ \text{constant } c_r, \\ (ii)\ \nabla_p H(x,p) \text{ exists and is Lipschitz continuous in } x \text{ on } B(0,r), \text{ uniformly} \\ \quad\ \text{for } p\in S^{n-1}. \end{cases}$$

We recall that (H) was introduced for the minimum time problem in [10] to derive sufficient conditions for the semiconcavity of the minimum time function. We refer the reader to [14, 6, 7] for a detailed discussion of (H).

Remark 1. Actually, in [6, 14] the authors suppose that the Hamiltonian $H^+(x,p) := \sup_{v\in F(x)}\langle v,p\rangle$ satisfies (H). On the other hand, it is easy to compute that $H^+(x,p) = H(x,-p)$, and so H^+ satisfies (H) if and only if so does H.

We recall below a classical result known as *Maximum principle* for the minimum time problem. It yields as necessary condition for the optimality of a trajectory $x(\cdot)$ the existence of a *dual arc* $p(\cdot)$ such that the pair (x,p) satisfies an Hamiltonian inclusion and a transversality condition.

Theorem 1. *Assume that (SH) and (H) hold. Suppose $x(\cdot)$ is an optimal solution of the minimum time problem $\mathscr{P}(x)$, reaching the target $\mathscr{K}$ at time $T := T(x)$. Then there exists an absolutely continuous arc $p:[0,T]\to\mathbb{R}^n$, $p(\cdot)\neq 0$, such that for a.e. $t\in[0,T]$,*

$$\begin{cases} -\dot{x}(t) = \nabla_p H(x(t),p(t)), \\ \dot{p}(t) \in \partial_x^- H(x(t),p(t)), \end{cases} \qquad p(T)\in N^C_{\mathscr{K}}(x(T)). \tag{10}$$

The classical formulation of the above theorem (see, for instance, [16]) is expressed in terms of the "complete" Hamiltonian system $(\dot{x},\dot{p})\in\partial H(x,p)$ (where ∂H stays for Clarke's generalized gradient of H in (x,p)). However, the "splitting Lemma" in [7] (Lemma 2.9) guarantees that under our assumptions these two formulations are equivalent.

Remark 2. Let us give two remarks.

(a) Let (x,p) be a solution to the Hamiltonian inclusion

$$\begin{cases} -\dot{x}(t) \in \partial_p^- H(x(t),p(t)), \\ \dot{p}(t) \in \partial_x^- H(x(t),p(t)), \end{cases} \qquad \text{a.e. in } [t_0,T]. \tag{11}$$

Then, there are only two possible cases:

- either $p(t)\neq 0$ for all $t\in[t_0,T]$,
- or $p(t)=0$ for all $t\in[t_0,T]$.

Indeed, consider $r > 0$ such that $x([t_0,T]) \subset B(0,r)$. If we denote by c_r a Lipschitz constant for F on $B(0,r)$, then $c_r|p|$ is a Lipschitz constant for $H(\cdot,p)$ on $B(0,r)$. Thus,

$$|\zeta| \le c_r|p| \quad \forall \zeta \in \partial_x^- H(x,p),\ \forall x \in B(0,r),\ \forall p \in \mathbb{R}^n.$$

Hence, $|\dot{p}(s)| \le c_r|p(s)|$ for a.e. $s \in [t_0,T]$. Therefore, Gronwall's Lemma allows to conclude.

(b) If (x,p) is a solution to (11), then for any $\lambda > 0$ the pair $(x,\lambda p)$ solves (11) as well. Indeed, by the positive 1-homogeneity in p of the Hamiltonian, that is $H(x,\lambda p) = \lambda H(x,p)$ for all $\lambda > 0$, x, $p \in \mathbb{R}^n$, it follows that $\partial_x H(x,\lambda p) = \lambda \partial_x H(x,p)$ and $\partial_p H(x,\lambda p) = \partial_p H(x,p)$ for all $\lambda > 0$, x, $p \in \mathbb{R}^n$. Thus, the proof of our claim is an easy verification.

For our aims, sometimes we shall need more refined necessary conditions than the ones in Theorem 1. Assuming the interior sphere property on the target $\mathscr{K}$ allows to further specify the transversality condition.

Proposition 2. *Assume that (SH) and (H) hold. Suppose $x(\cdot)$ is an optimal solution for the minimum time problem $\mathscr{P}(x)$, reaching the target $\mathscr{K}$ at time $T := T(x)$, and that there exists $0 \neq \nu \in N^P_{\mathscr{K}^C}(x(T))$ realized by a ball of radius R, that is,*

$$B\left(x(T) + R\frac{\nu}{|\nu|}, R\right) \subset \mathscr{K}.$$

Then there exists an absolutely continuous arc $p : [0,T] \to \mathbb{R}^n$, $p(\cdot) \neq 0$, such that for a.e. $t \in [0,T]$,

$$\begin{cases} -\dot{x}(t) = \nabla_p H(x(t),p(t)), \\ \dot{p}(t) \in \partial_x H(x(t),p(t)), \end{cases} \qquad p(T) = -\nu. \tag{12}$$

Proof. The trajectory $x(\cdot)$ is time-optimal even for the problem obtained replacing the target $\mathscr{K}$ by the ball $B_1 := B\big(x(T) + R\nu\,|\,\nu\,|^{-1}, R\big)$. Moreover, $N^C_{B_1}(x(T)) = \{-\nu\}$. Thus, applying Theorem 1 to this new problem we prove our claim.

Remark 3. If we suppose in addition that

$$\mu(-\nu) := H(x(T),-\nu)^{-1} > 0, \tag{13}$$

then the above theorem together with Remark 2 (b) gives that there exists an absolutely continuous arc $p : [0,T] \to \mathbb{R}^n$, $p(\cdot) \neq 0$, such that (x,p) solves, for a.e. $t \in [0,T]$,

$$\begin{cases} -\dot{x}(t) = \nabla_p H(x(t),p(t)), \\ \dot{p}(t) \in \partial_x H(x(t),p(t)), \end{cases} \qquad p(T) = -\mu(-\nu)\nu. \tag{14}$$

In addition to our assumptions on F and H, further hypotheses on the target set $\mathscr{K}$ might be needed, such as the inner sphere property and the so-called *Petrov condition* we recall below:

(PC) $\exists \delta > 0$ such that $H(x,\zeta) \ge \delta\|\zeta\|$ for all $x \in \partial\mathscr{K}$ and all $\zeta \in N^P_{\mathscr{K}}(x)$.

Assumption (PC) turns out to be equivalent to the Lipschitz continuity of the minimum time function $T(\cdot)$ in a neighborhood of $\mathscr{K}$. It is also necessary for the semiconcavity of $T(\cdot)$ up to a boundary of $\mathscr{K}$ and equivalent to the validity of a bound of T in terms of the distance function from the target $\mathscr{K}$, which is defined as

$$d_{\mathscr{K}} : \mathbb{R}^n \to \mathbb{R}^+, \quad d_{\mathscr{K}}(x) := \inf\{|\, y - x \,| : y \in \mathscr{K}\}.$$

Recall, among the other things, that assuming Petrov's condition on the target $\mathscr{K}$ guarantees that (13) always holds true. For a comprehensive treatment and further references on this subject we refer to the book [13]. In sections 3 and 4.3 we shall also assume that $\mathscr{K}$ is the closure of its interior and

(A) $\partial\mathscr{K}$ is an $(n-1)$-dimensional manifold of class C^2.

Whenever (A) holds true, $\mathscr{K}$ satisfies the inner sphere property with a uniform positive radius. Moreover, the signed distance from the target $\mathscr{K}$, that is,

$$b_{\mathscr{K}} : \mathbb{R}^n \to \mathbb{R}, \quad b_{\mathscr{K}}(\cdot) := d_{\mathscr{K}}(\cdot) - d_{\mathscr{K}^C}(\cdot),$$

is a function of class C^2 in a neighborhood of $\partial\mathscr{K}$, and $-\nabla b_{\mathscr{K}}(\xi)$ is a proximal inner normal to $\mathscr{K}$ at $\xi \in \partial\mathscr{K}$ with unit norm.

3 Conjugate times for the minimum time problem

The aim of this section is to extend the main result in [15] to the minimum time problem. More precisely, we show that the absence of conjugate times is equivalent to the propagation of the local regularity of the minimum time function. Let us mention that a partial result in this framework has been recently given in [18]. On the other hand, our notion of conjugate time is more in the spirit of [21] and allows to recover a stronger result than the one in [18].

3.1 Conjugate times for the minimum time problem

In this section, we assume (SH), (PC), and (A) and suppose that the Hamiltonian H is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$. Given $\xi \in \partial\mathscr{K}$, set $g(\xi) := \mu(\xi)\nabla b_{\mathscr{K}}(\xi)$, where $\mu(\xi)$ is the positive constant $H(\xi, \nabla b_{\mathscr{K}}(\xi))^{-1}$. Recall also that, thanks to (A), the function g is of class C^1 in a neighborhood of $\partial\mathscr{K}$. Therefore, we denote by $(Y(\xi,\cdot), P(\xi,\cdot))$ (or, briefly, by $(Y(\cdot), P(\cdot))$) the solution of the *backward Hamiltonian system*

$$\begin{cases} \dot{Y}(t) = \nabla_p H(Y(t), P(t)), & Y(0) = \xi, \\ -\dot{P}(t) = \nabla_x H(Y(t), P(t)), & P(0) = g(\xi). \end{cases} \tag{15}$$

We recall that for any $\xi \in \partial\mathscr{K}$ the solution $(Y(\cdot),P(\cdot))$ to (15) is well-defined on $[0,+\infty)$ and the functions Y, P are of class C^1 with respect to ξ and the time in $\partial\mathscr{K}\times[0,+\infty)$ (for the proof of these facts see, for instance, Section 3 in [21]).
Since $\partial\mathscr{K}$ is a C^2-manifold of dimension $n-1$, for any $\xi_0 \in \partial\mathscr{K}$ there exist a C^2 local parameterization of $\partial\mathscr{K}$:

$$\phi : A \subset \mathbb{R}^{n-1} \to \mathbb{R}^n, \quad \eta \to \phi(\eta) = \xi.$$

Set $\eta_0 := \phi^{-1}(\xi_0)$. Let us denote by $Y_{\xi,t}(\xi,t)$ and $P_{\xi,t}(\xi,t)$ the Jacobians of $Y(\phi(\cdot),\cdot)$ and $P(\phi(\cdot),\cdot)$ with respect to the state variable $\eta \in \mathbb{R}^{n-1}$ and time, that is,

$$Y_{\xi,t}(\xi,t) = Y_{\eta,s}(\phi(\eta),t), \quad P_{\xi,t}(\xi,t) = P_{\eta,s}(\phi(\eta),t).$$

Therefore, note that $Y_{\xi,t}(\xi,t)$ and $P_{\xi,t}(\xi,t)$ belong to $\mathbb{R}^{n\times n}$ and the pair $(Y_{\xi,t},P_{\xi,t})$ solves the *variational system*

$$\begin{cases} \dot{Y}_{\xi,t} = H_{xp}(Y,P)Y_{\xi,t} + H_{pp}(Y,P)P_{\xi,t},\ Y_{\xi,t}(\xi,0) = \left(\frac{\partial\phi}{\partial\eta}(\eta), \nabla_p H(\xi,p)\right), \\ -\dot{P}_{\xi,t} = H_{xx}(Y,P)Y_{\xi,t} + H_{px}(Y,P)P_{\xi,t},\ P_{\xi,t}(\xi,0) = \left(\frac{\partial g}{\partial\eta}(\phi(\eta)), -\nabla_x H(\xi,p)\right), \end{cases} \tag{16}$$

where we have set $p := \mu(\xi)\nabla b_{\mathscr{K}}(\xi)$. Matrix $Y_{\xi,t}(\xi,0)$ is invertible; indeed, by (PC) and the choice of $\mu(\cdot)$ it follows that

$$0 \neq \mu(\xi)^{-1}H(\xi,\mu(\xi)\nabla b_{\mathscr{K}}(\xi)) = \langle \nabla_p H(\xi,\mu(\xi)\nabla b_{\mathscr{K}}(\xi)), \nabla b_{\mathscr{K}}(\xi)\rangle.$$

Thus, the vector $\mu(\xi)\nabla b_{\mathscr{K}}(\xi)$ is non-characteristic for the data $g(\cdot)$, that is,

$$\langle \nabla_p H(\xi,\mu(\xi)\nabla b_{\mathscr{K}}(\xi)), \nabla b_{\mathscr{K}}(\xi)\rangle \neq 0.$$

It is natural to introduce the following definition of *conjugate time*.

Definition 1. Let $\xi_0 \in \partial\mathscr{K}$ and let ϕ a local C^2 parameterization of $\partial\mathscr{K}$ near ξ_0. Let $(Y_{\xi,t},P_{\xi,t})$ be the solution to (16). Define

$$\bar{t} = \sup\{t \in [0,+\infty) :\ \det Y_{\xi,t}(\xi_0,s) \neq 0 \ \text{ for all } s \in [0,t]\}.$$

The time $\bar{t}$ is called *conjugate* for ξ_0 if $\bar{t} < +\infty$.

Thus, if $\bar{t}$ is conjugate for ξ_0 then $\det Y_{\xi,t}(\xi_0,\bar{t}) = 0$.
Note that the solution $(Y_{\xi,t},P_{\xi,t})$ to (16) depends on the parameterization ϕ. On the other hand, the ranks of the values of the maps $Y_{\xi,t}(\xi_0,\cdot)$ and $P_{\xi,t}(\xi_0,\cdot)$ are independent of the particular choice of ϕ, as well as the above definition of conjugate time.
By standard techniques one deduces that if $\det Y_{\xi,t}(\xi_0,t) \neq 0$, then there exists a neighborhood of (ξ_0,t) in $\partial\mathscr{K}\times\mathbb{R}$ such that the matrix $Y_{\xi,t}(\xi,s)$ is nonsingular for any vector (ξ,s) in such a neighborhood. Furthermore, if there are no conjugate times for ξ_0 on some interval $[0,a]$, then the map $Y(\cdot,\cdot)$ provides a diffeomorphism from a neighborhood $J_{\xi_0}\times U_t$ of (ξ_0,t) in $\partial\mathscr{K}\times\mathbb{R}$ onto its image for all $t \in [0,a]$.

Moreover, it is easy to check that the function $R(\xi_0,t) := P_{\xi,t}(\xi_0,t)Y_{\xi,t}^{-1}(\xi_0,t)$, as long as $Y_{\xi,t}(\xi_0,t)$ is invertible, solves the Riccati equation

$$\begin{cases} \dot{R} + H_{px}(Y,P)R + RH_{xp}(Y,P) + RH_{pp}(Y,P)R + H_{xx}(Y,P) = 0, \\ R(\xi_0,0) = P_{\xi,t}(\xi_0,0)Y_{\xi,t}^{-1}(\xi_0,0). \end{cases} \tag{17}$$

For a fixed $\theta \in \mathbb{R}^n \setminus \{0\}$ and for any $t > 0$, let us denote by $w(t)$ the $2n$-vector given by $(Y_{\xi,t}(\xi_0,t)\theta, P_{\xi,t}(\xi_0,t)\theta)$. It is easy to check that $w(\cdot)$ solves a linear differential system with nonzero initial data, since $Y_{\xi,t}(\xi_0,0)$ has rank n. By well-known properties of linear systems, it follows that $w(t) \neq 0$ for all $t > 0$. This means that for any $\theta \in \mathbb{R}^n \setminus \{0\}$ and $t > 0$,

$$Y_{\xi,t}(\xi_0,t)\theta = 0 \;\Rightarrow\; P_{\xi,t}(\xi_0,t)\theta \neq 0.$$

Therefore, it is easy to understand that a time $\bar{t}$ is conjugate for ξ_0 if and only if $[0,\bar{t})$ is the maximal interval of existence of the solution $R(\xi_0,\cdot)$ to (17) and $\bar{t} < +\infty$. Thus, $\bar{t}$ is a finite blow-up time for $R(\xi_0,\cdot)$, that is,

$$\lim_{t \nearrow \bar{t}} \| R(\xi_0,t) \| = +\infty.$$

If (Y,P) is given on a finite time interval $[0,T]$, then the above definition of conjugate time can be adapted, by saying that $\bar{t} \in [0,T]$ is a conjugate time for ξ_0 if and only if $\det Y_{\xi,t}(\xi_0,t) \neq 0$ for all $t \in [0,\bar{t})$ and $\det Y_{\xi,t}(\xi_0,\bar{t}) = 0$. Equivalently, $\bar{t}$ is a conjugate time for ξ_0 if and only if $R_{\xi,t}(\xi_0,\cdot)$ is well defined on $[0,\bar{t})$ and $\lim_{t \nearrow \bar{t}} \| R_{\xi,t}(\xi_0,t) \| = +\infty$.

3.2 Local regularity of the minimum time function and conjugate times

Let $\xi_0 \in \partial\mathscr{K}$ and $t \geq 0$. Given an open neighborhood $V_{\xi_0} \times I_t$ of (ξ_0,t) in $\partial\mathscr{K} \times \mathbb{R}$, define the set

$$M(V_{\xi_0} \times U_t) := \{(Y(\xi,s),P(\xi,s)) |\ (Y,P) \text{ solves (15) with } \xi \in V_{\xi_0},\ s \in U_t\}.$$

When $t > 0$, the set U_t may be viewed as an interval of the form $(t-b,t+b)$ for some $b > 0$ and when $t = 0$ as the interval $(0,b)$.

Theorem 2. *Let us assume (SH), (PC) and (A) and suppose that the Hamiltonian H is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$. Fix $\bar{t} > 0$. Then, the following two statements are equivalent:*

(i) for all $t \in [0,\bar{t}]$, there exists an open neighborhood $V_{\xi_0} \times I_t$ of (ξ_0,t) in $\partial\mathscr{K} \times \mathbb{R}$ such that the set

$$\mathscr{D}(V_{\xi_0} \times I_t) := \{Y(\xi,s) | (Y,P) \textit{ solves } (15), \textit{ with } \xi \in V_{\xi_0},\ s \in I_t\} \tag{18}$$

is an open subset of $\mathbb{R}^n$*, and* $M(V_{\xi_0} \times I_t)$ *is the graph of a* C^1 *function on* $\mathscr{D}(V_{\xi_0} \times U_t)$*;*

(ii) there are no conjugate times for ξ_0 *on* $[0,\bar{t}]$.

In [15], Caroff and Frankowska analysed the link between conjugate points and regularity of the value function V for Bolza optimal control problems, showing that the first emergence of a conjugate point corresponds to the first time when V stops to be locally smooth along optimal trajectories. In Theorem 2 we prove that the same kind of result holds true also for the minimum time problem. We note that our result cannot be deduced from the one in [15]—even though the technique of proof is similar—because the definition of conjugate time we use in this paper is different from the one therein. For this reason, we give below the proof of the implication $(ii) \Rightarrow (i)$, which is the one needed to derive Theorem 6 below.

Proof. Suppose that there are no conjugate times for ξ_0 on $[0,\bar{t}]$. We want to show that there exists a neighborhood $V_{\xi_0} \times I_t$ of (ξ_0,t) in $\partial\mathscr{K} \times \mathbb{R}$ such that $M(V_{\xi_0} \times U_t)$ is a graph of a C^1 function on $\mathscr{D}(V_{\xi_0} \times I_t)$, for all $t \in [0,\bar{t}]$. Actually, we shall prove, first, that $M(V_{\xi_0} \times U_t)$ is a graph of a Lipschitz function with Lipschitz constant uniform in $[0,\bar{t}]$. So, proceeding by contradiction, let us fix any neighborhood $V_{\xi_0} \times I_t$ of (ξ_0,t) in $\partial\mathscr{K} \times \mathbb{R}$ and let us consider the compact set $\Pi_t := \overline{M(V_{\xi_0} \times I_t)}$ for all $t > 0$. It is a well-known fact that there exists a time $t^* > 0$ such that Π_t is a graph of a Lipschitz function for all $t \in [0,t^*]$. Let $a = \sup \mathscr{T}$, where

$$\begin{aligned}\mathscr{T} := \{t \in [0,\bar{t}] :\ & \exists k_t \geq 0 \text{ s.t. } \Pi_s \text{ is a graph of a } k_t\text{-Lipschitz function} \\ & \Phi_s : \overline{\mathscr{D}(V_{\xi_0} \times I_s)} \to \mathbb{R}^n \text{ for all } s \in [0,t]\}.\end{aligned}$$

Aiming to a contradiction, suppose that $a \notin \mathscr{T}$, i.e., Π_a is not the graph of a k-Lipschitz function. Then, fix $t \in [0,a)$. Since $\det Y_{\xi,t}(\xi_0,t) \neq 0$, without loss of generality, we can suppose that for any vector $(\xi,s) \in \overline{V_{\xi_0}} \times \overline{U_t}$ we have that $\det Y_{\xi,t}(\xi,s) \neq 0$. Moreover, $Y(\cdot,\cdot) : V_{\xi_0} \times U_t \to \mathbb{R}^n$ is an injective continuous map. Thus, $Y(V_{\xi_0},U_t)$ is an open set by Brouwer's invariance of domain theorem. Note that $\mathscr{D}(V_{\xi_0} \times U_t) \equiv Y(V_{\xi_0},U_t)$. Consequently, $\mathscr{D}(V_{\xi_0} \times U_t)$ is open and its closure is

$$\overline{\mathscr{D}(V_{\xi_0} \times U_t)} := \{Y(\xi,s) | (Y,P) \text{ solves (15) with } \xi \in \overline{V_{\xi_0}},\ s \in \overline{U_t}\}.$$

Note that the map Φ_t is a.e. differentiable on $\mathscr{D}(V_{\xi_0} \times U_t)$ for all $t \in [0,a)$. Since Π_a is not a Lipschitz graph, there exist two sequences $t_i \nearrow a$ and $\{x_i\}_{i\in\mathbb{N}} \subset \mathscr{D}(V_{\xi_0} \times U_{t_i})$ such that

$$\| D\Phi_{t_i}(x_i) \| \to +\infty.$$

Equivalently, we can find a sequence of vectors $\{u_i,v_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^n \times \mathbb{R}^n$ such that $D\Phi_{t_i}(x_i)u_i = v_i$, $|v_i| = 1$ for all $i \in \mathbb{N}$ and $|u_i| \to 0$ as $i \to +\infty$.

Since $\{x_i\}_{i\in\mathbb{N}} \subset \mathscr{D}(V_{\xi_0} \times U_{t_i})$, there exist a sequence of vectors $\{\xi_i\}_i \subset V_{\xi_0}$ and one of times $s_i \in U_{t_i}$ such that the solution $(Y(\xi_i,\cdot),P(\xi_i,\cdot))$ to (15) solves $Y(\xi_i,s_i) = x_i$ and $P(\xi_i,s_i) = \Phi_{t_i}(x_i)$. Now, let us consider the linearization of the system associated to $(Y(\xi_i,\cdot),P(\xi_i,\cdot))$ at (x_i,p_i) given by the solution (w_i,q_i) to

$$\begin{cases} \dot{w}_i = H_{xp}(Y(\xi_i,t),P(\xi_i,t))w_i + H_{pp}(Y(\xi_i,t),P(\xi_i,t))q_i, \; w_i(s_i) = u_i, \\ -\dot{q}_i = H_{xx}(Y(\xi_i,t),P(\xi_i,t))w_i + H_{px}(Y(\xi_i,t),P(\xi_i,t))q_i, \;\; q_i(s_i) = v_i. \end{cases} \tag{19}$$

Consequently,

$$D\Phi_{t_i}(x_i)w_i(s_i) = q_i(s_i), \;\; |w_i(s_i)| \to 0 \text{ and } |q_i(s_i)| = 1. \tag{20}$$

After possibly passing to a subsequence, we may assume that the sequence $\{\xi_i\}_{i\in\mathbb{N}}$ converges to some vector $\overline{\xi} \in \overline{V_{\xi_0}}$ and $\{s_i\}_{i\in\mathbb{N}}$ to some time $\bar{s} \in \overline{U_t}$, as $i \to +\infty$. Then, passing to the limit as $i \to +\infty$, it is easy to deduce from (20) that the vector $\overline{\xi}$ has a conjugate time equal to $\bar{s}$, i.e., $\det Y_{\xi,t}(\overline{\xi},\bar{s}) = 0$. Since $(\overline{\xi},\bar{s}) \in \overline{V_{\xi_0}} \times \overline{U_t}$, we obtain a contradiction. Therefore, Π_t is a graph of a Lipschitz function for all $t \in [0,\bar{t}]$. Since g is of class C^1 in a neighborhood of $\partial\mathscr{K}$, by well-known properties of linearized systems we deduce that, for every parameterization ϕ of V_{ξ_0}, $\Phi_t \circ \phi^{-1}$ is of class C^1 and (i) holds true.

Remark 4. Suppose that the map Φ_t is of class C^1 on the set $\mathscr{D}(V_{\xi_0} \times U_t)$ for all $t \in [0,\bar{t}]$. Then, it is easy to understand that its Jacobian is given by: for all $\xi \in V_{\xi_0}$ and $s \in U_t$,

$$D\Phi_t(Y(\xi,s)) = P_{\xi,t}(\xi,s)Y_{\xi,t}(\xi,s)^{-1},$$

in the sense that the matrix

$$P_{\eta,t}(\phi(\eta),s)Y_{\eta,t}(\phi(\eta),s)^{-1}$$

represents the Jacobian of Φ_t at $Y(\xi,s)$ in the system of local coordinates $(\eta_1,...\eta_{n-1})$ induced by a parameterization ϕ of V_{ξ_0}.

A characteristic $Y(\xi,\cdot)$, with $\xi \in \partial\mathscr{K}$, is said to be *optimal* in some interval $[0,\tau]$ if it coincides with an optimal trajectory $y(\cdot)$ starting from $Y(\xi,\tau)$ running backward in time, that is, $Y(\xi,t) = y(\tau - t)$, for all $t \in [0,\tau]$. By the classical method of characteristics, one can deduce that any characteristic $Y(\xi,\cdot)$ is optimal in $[0,\tau^*)$ for some time $\tau^* > 0$. Theorem 2 allows to deduce that this result holds true as long as there are no conjugate times.

Corollary 1. *Let us assume (A) and suppose that H is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \times \{0\}))$. If there are no conjugate times for ξ_0 on the interval $[0,\bar{t}]$, then there exists a neighborhood of ξ_0 in $\partial\mathscr{K}$, V_{ξ_0}, such that $Y(\xi,\cdot)$ is optimal on $[0,\bar{t}]$ for any $\xi \in V_{\xi_0}$.*

3.3 A characterization of conjugate times

In this subsection, let us assume (SH), (PC) and (A) and suppose that the Hamiltonian H is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$. Let us denote by $Y_\xi(\xi,t)$ and $P_\xi(\xi,t)$ the Jacobian of $Y(\phi(\cdot),t)$ and $P(\phi(\cdot),t)$ with respect to the state variable $\eta \in \mathbb{R}^{n-1}$ evaluated at (ξ,t), that is,

$$Y_\xi(\xi,t) = Y_\eta(\phi(\eta),t), \quad P_\xi(\xi,t) = P_\eta(\phi(\eta),t).$$

One can easily check that the pair $(Y_\xi(\xi,\cdot),\ P_\xi(\xi,\cdot))$ takes values in $\mathbb{R}^{n\times(n-1)} \times \mathbb{R}^{n\times(n-1)}$ and solves the system

$$\begin{cases} \dot{Y}_\xi = H_{xp}(Y,P)Y_\xi + H_{pp}(Y,P)P_\xi,\ Y_\xi(\xi,0) = \frac{\partial\phi}{\partial\eta}(\eta), \\ -\dot{P}_\xi = H_{xx}(Y,P)Y_\xi + H_{px}(Y,P)P_\xi,\ P_\xi(\xi,0) = \frac{\partial g}{\partial\eta}(\phi(\eta)). \end{cases} \tag{21}$$

In the case of a strictly convex Hamiltonian in p, the notion of conjugate time can be characterized through the solution of the above system (see [21, Theorem 6.1]). Let us now introduce the hypothesis:

(H2) The kernel of $H_{pp}(x,p)$ has dimension equal to 1 for every $(x,p) \in \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})$, i.e., $\ker H_{pp}(x,p) = p\,\mathbb{R}$.

We will show that also under the weaker assumption $(H2)$ a similar characterization can be provided.

Proposition 3. *Suppose that H satisfies $(H2)$. For any $t > 0$, it holds that* $\det Y_{\xi,t}(\xi,t) = 0$ *if and only if* $rkY_\xi(\xi,t) < n-1$.

To prove the above proposition, we need the following lemma.

Lemma 1. *Under the assumptions of Proposition 3, it holds that, for any $t > 0$,*

$$\frac{d}{ds}\det Y_{\xi,s}(\xi,s)\,|_{s=t} \neq 0 \quad \Longleftrightarrow \quad rkY_{\xi,t}(\xi,t) = n-1.$$

Proof. First, suppose that $rkY_{\xi,t}(\xi,t) = n-1$. Following the same reasoning as in the proof of [21, Lemma 4.3], we have that

$$\frac{d}{ds}\det Y_{\xi,s}(\xi,s)\,|_{s=t} = tr\left(H_{pp}(Y(\xi,t),P(\xi,t))P_{\xi,t}(\xi,t)Y^+_{\xi,t}(\xi,t)\right),$$

where A^+ denotes the transpose of the matrix of the cofactors of a matrix A, that is, $AA^+ = A^+A = (detA)I_n$. Moreover, if θ is such that $\ker Y_{\xi,t}(\xi_0,t) = \theta\mathbb{R}$, then by [21, Lemma 4.2] there exists $c > 0$ such that

$$\frac{d}{ds}\det Y_{\xi,s}(\xi,s)\,|_{s=t} = cH_{pp}(Y(\xi,t),P(\xi,t))P_{\xi,t}(\xi,t)\theta \cdot P_{\xi,t}(\xi,t)\theta. \tag{22}$$

We claim that $P_{\xi,t}(\xi,t)\theta \notin \ker H_{pp}(Y(\xi,t),P(\xi,t))$. If not, there exists $\lambda \in \mathbb{R}\setminus\{0\}$ such that $P_{\xi,t}(\xi,t)\theta = \lambda P(\xi,t)$. Now, observe that for all $\xi \in \partial\mathscr{K}$ and all $t \geq 0$ it holds that $H(Y(\xi,t),P(\xi,t)) = 1$. Hence, taking the Jacobian of this map at (ξ,t) and recalling that $\ker Y_{\xi,t}(\xi_0,t) = \theta\mathbb{R}$ we obtain that

$$\begin{aligned} 0 &= \langle Y_{\xi,t}(\xi,t)\theta, H_x(Y(\xi,t),P(\xi,t))\rangle + \langle P_{\xi,t}(\xi,t)\theta, H_p(Y(\xi,t),P(\xi,t))\rangle \\ &= \langle P_{\xi,t}(\xi,t)\theta, H_p(Y(\xi,t),P(\xi,t))\rangle. \end{aligned}$$

On the other hand, since we are assuming that $P_{\xi,t}(\xi,t)\theta = \lambda P(\xi,t)$, we have

$$\begin{aligned}&\langle P_{\xi,t}(\xi,t)\theta, H_p(Y(\xi,t),P(\xi,t))\rangle = \langle \lambda P(\xi,t)\theta, H_p(Y(\xi,t),P(\xi,t))\rangle\\ &= \lambda\langle P(\xi,t)\theta, H_p(Y(\xi,t),P(\xi,t))\rangle = \lambda H(Y(\xi,t),P(\xi,t)) = \lambda,\end{aligned}$$

that is in clear contradiction with the equality that is above it. This finally shows that $P_{\xi,t}(\xi,t_0)\theta \notin \ker H_{pp}(Y(\xi,t),P(\xi,t))$, and so from (22) we obtain that

$$\frac{d}{ds}\det Y_{\xi,s}(\xi,s)\mid_{s=t} > 0.$$

For the proof of the other implication, we refer the reader to the proof of [21, Lemma 4.3].

Proof (Proof of Proposition 3). It is sufficient to show that if $\det Y_{\xi,t}(\xi,t) = 0$ then $rkY_\xi(\xi,t) < n-1$. Aiming for a contradiction, suppose $\det Y_{\xi,t}(\xi,t) = 0$ but $rkY_\xi(\xi,t) = n-1$. Hence, the vectors $Y_{\eta_i}(\xi,t)$, $i = 1,\dots,n-1$, are linearly independent and, by continuity, there exists $\delta > 0$ such that for any time $s \in (t-\delta, t+\delta)$ the vectors $Y_{\eta_i}(\xi_0,t)$, $i = 1,\dots,n-1$, are still linearly independent. We can distinguish to cases:

1. there exists a sequence of times $t_k \to t$ as $k \to \infty$ such that $\det Y_{\xi,t}(\xi,t_k) = 0$ for all k,
2. there exists a constant $\delta' \in (0,\delta)$ such that $\det Y_{\xi,t}(\xi,s) \neq 0$ for all $s \in (t-\delta', t+\delta')$.

For the discussion of the first case, we refer the reader to the proof of [21, Theorem 6.1]. In the second case, we have that

$$\frac{d}{ds}\det Y_{\xi,s}(\xi,s) \neq 0 \quad \text{for all } s \in (t-\delta', t+\delta').$$

Then, Lemma 1 implies that $rkY_{\xi,t}(\xi,t) < n-1$ and this yields the contradiction.

Under the additional assumption $(H2)$, the above proposition gives an equivalent characterization of conjugate times considering only the spatial Jacobian of the map $Y(\cdot,\cdot)$. More specifically, it follows that a time $\bar{t}$ is conjugate for ξ_0 if and only if

$$\bar{t} = \sup\{t \in [0,+\infty) : \ rkY_\xi(\xi_0,s) = n-1 \text{ for all } s \in [0,t]\},$$

and $\bar{t} < +\infty$. Consequently, $rkY_\xi(\xi_0,\bar{t}) < n-1$.

Remark 5. Let us suppose that $(H2)$ holds true. If there is no conjugate time for ξ_0 on $[0,\bar{t}]$, then $Y(t,\cdot)$ maps a neighborhood I_{ξ_0} of ξ_0 in $\partial\mathscr{K}$ onto the level sets of the minimum time function, that is, for all $t \in [0,\bar{t}]$,

$$Y(t,\cdot) : I_{\xi_0} \to \Gamma_t \subset \mathbb{R}^{n-1},$$

where $\Gamma_t := \{x \in \mathbb{R}^n : \ T(x) = t\}$. Moreover, $Y(t,\cdot)$ gives a diffemorphism from a neighborhood of ξ_0 in $\partial\mathscr{K}$ onto an open neighborhood of Γ_t, for any time t smaller than the conjugate time $\bar{t}$.

4 First-order sensitivity relations for the minimum time problem and some applications

The scope of this section is twofold. First, we discuss some sensitivity relations of first order. Subsequently, we apply these results to derive sufficient conditions for the propagation of the regularity of the minimum time function along optimal trajectories.

4.1 Proximal subdifferentiability of the minimum time function

The forward propagation of the dual arc into the proximal subdifferential of T is already known for minimum time problems when the dynamic is described by a control system with sufficiently smooth dynamics (see [18]). We shall extend this result to the differential inclusion case.

Theorem 3. *Assume (SH) and (H) hold. Let $x_0 \in \mathscr{R} \setminus \mathscr{K}$, let $\bar{x}$ be an optimal trajectory starting from x_0 reaching $\mathscr{K}$ in time $T(x_0)$, and $\overline{p} : [0,T(x_0)] \to \mathbb{R}^n$ be an arc such that $(\bar{x},\overline{p})$ solves the system: for a.e. $t \in [0,T(x_0)]$,*

$$\begin{cases} -\dot{x}(t) \in \partial_p^- H(x(t),p(t)), & x(0) = x_0 \\ \dot{p}(t) \in \partial_x^- H(x(t),p(t)), & p(0) \in \partial^{-,P} T(x_0). \end{cases}$$

Then, there exist constants $c, r > 0$ such that, for all $t \in [0,T(x_0))$ and $h \in B(0,r)$,

$$T(\bar{x}(t)+h) - T(\bar{x}(t)) \geq \ \langle \overline{p}(t), h\rangle - c \mid h \mid^2 .$$

Consequently,

$$\overline{p}(t) \in \partial^{-,P} T(\bar{x}(t)) \quad \textit{for all } t \in [0,T(x_0)). \tag{23}$$

Proof. First of all, recall that $0 \notin \partial^{-,P} T(x_0)$ (see, for instance, [25, Theorem 5.1]), and so the dual arc $\overline{p}$ never vanishes on $[0,T(x_0)]$ by Remark 2. Since $\overline{p}(0) \in \partial^{-,P} T(x_0)$, there exist $c_0, r_0 > 0$ such that for every $h \in B(0,r_0)$,

$$T(x_0+h) - T(x_0) \geq \langle \overline{p}(0), h\rangle - c_0|h|^2. \tag{24}$$

Fix $t \in (0,T(x_0))$. Recall that $\bar{x}(\cdot)$ is the unique solution of the final value problem

$$\begin{cases} -\dot{x}(s) = \nabla_p H(x(s),\overline{p}(s)) & \text{for all } s \in [0,t]\,, \\ x(t) = \bar{x}(t). \end{cases}$$

For all $h \in B$, let $x_h(\cdot)$ be the solution of the system

$$\begin{cases} -\dot{x}(s) = \nabla_p H(x(s), \overline{p}(s)) & \text{for all } s \in [0,t]\,, \\ x(t) = \overline{x}(t) + h. \end{cases}$$

From the optimality of $\overline{x}(\cdot)$ and the dynamic programming principle we deduce that

$$\begin{aligned} T(\overline{x}(t)+h) - T(\overline{x}(t)) - \langle \overline{p}(t), h\rangle &= T(x_h(t)) - T(\overline{x}(t)) - \langle \overline{p}(t), h\rangle \\ &\geq T(x_h(0)) - T(x_0) - \langle \overline{p}(t), h\rangle. \end{aligned} \tag{25}$$

From the sublinearity of F and $(H)(ii)$, using a standard argument based on Gronwall's lemma, one can show that there exists $k > 0$, independent of $t \in (0, T(x_0))$, such that

$$\|x_h - \overline{x}\|_\infty \leq e^{kT} \mid h \mid, \quad \forall\, h \in B. \tag{26}$$

For all $h \in B(0,r)$ with $r := \min\{1, r_0 e^{-kT}\}$, by (24), (25) and (26) we have that

$$T(\overline{x}(t)+h) - T(\overline{x}(t)) - \langle \overline{p}(t), h\rangle \geq -\langle \overline{p}(t), h\rangle + \langle \overline{p}(0), x_h(0) - x_0\rangle - c_0 \mid x_h(0) - x_0 \mid^2. \tag{27}$$

Moreover,

$$\begin{aligned} \langle \overline{p}(0), x_h(0) - x_0\rangle - \langle \overline{p}(t), h\rangle &= -\int_0^t \frac{d}{ds}\langle \overline{p}(s), x_h(s) - \overline{x}(s)\rangle\, ds \\ &= -\int_0^t \Big(\langle \dot{\overline{p}}(s), x_h(s) - \overline{x}(s)\rangle + \langle \overline{p}(s), \dot{x}_h(s) - \dot{\overline{x}}(s)\rangle \Big) ds \\ &= -\int_0^t \Big(\langle \dot{\overline{p}}(s), x_h(s) - \overline{x}(s)\rangle - H(x_h(s), \overline{p}(s)) + H(\overline{x}(s), \overline{p}(s)) \Big) ds. \end{aligned}$$

Since $\dot{\overline{p}}(s) \in \partial_x^- H(\overline{x}(s), \overline{p}(s))$ a.e. in $[0, T(x_0)]$, assumption (H) (i) implies that

$$\langle \overline{p}(0), x_h(0) - x_0\rangle - \langle \overline{p}(t), h\rangle \geq c_2 \int_0^t \mid \overline{p}(s) \mid\mid x_h(s) - \overline{x}(s) \mid^2, \tag{28}$$

where c_2 is a suitable constant independent from $t \in (0, T(x_0))$. From (26)-(28) we obtain our conclusion.

Theorem 4. *Assume (SH) and (H) hold. Let $x_0 \in \mathscr{R} \setminus \mathscr{K}$, let $\overline{x}$ be an optimal trajectory starting from x_0 reaching $\mathscr{K}$ in time $T(x_0)$, and $\overline{p} : [0, T(x_0)] \to \mathbb{R}^n$ be an arc such that $(\overline{x}, \overline{p})$ solves the system: for a.e. $t \in [0, T(x_0)]$,*

$$\begin{cases} -\dot{x}(t) \in \partial_p^- H(x(t), p(t)), & x(0) = x_0 \\ \dot{p}(t) \in \partial_x^- H(x(t), p(t)), & p(0) \in \partial^- T(x_0). \end{cases} \tag{29}$$

Then, $\overline{p}(\cdot)$ satisfies

$$\overline{p}(t) \in \partial^- T(\overline{x}(t)) \quad \textit{for all } t \in [0, T(x_0)). \tag{30}$$

Proof. The proof of the case $\overline{p} \neq 0$ is similar to the proof of the above theorem. As the case $\overline{p} = 0$, we refer the reader to the proof of [6, Theorem 2.1], where we have described a strategy for constructing perturbations of the optimal trajectory when the dual arc is vanishing.

4.2 First application: differentiability of the minimum time function

Here, we provide some sufficient conditions in order that the optimal trajectories starting from a point in the domain of differentiability of the minimum time function $T(\cdot)$ stay in such a set whenever Petrov's condition is satisfied at the final point of the optimal trajectories. The same result has been obtained in [18] in the case of smooth targets and smooth control systems with a Hamiltonian of class $C^{1,1}$. Let us also mention that the fact that $T(\cdot)$ is differentiable along an optimal trajectory starting from x for all time in the open interval $(0, T(x))$ has been proved earlier in [8] in the case of exit-time problems with smooth control systems and a strongly convex Hamiltonian in p, under the Petrov condition (PC). More recently, in [11] this result was extended to the case of differential inclusions whit a strictly quasi-convex Hamiltonian in p. On the other hand, in our context the Hamiltonian is no longer strictly convex being 1-homogeneous in p and, in general, is not strictly quasi-convex, as shown by [11, Example 1.]. Here, the Hamiltonian is only assumed to satisfy (H), while the target is supposed to satisfy an interior sphere property.

Theorem 5. *Assume that (SH) and (H) hold. Let $x_0 \in \mathscr{R} \setminus \mathscr{K}$, let $\bar{x}$ be an optimal trajectory starting from x_0 reaching the target $\mathscr{K}$ at time $T(x_0)$. Let $R > 0$, and suppose that $\mathscr{K}$ is a nonempty subset of $\mathbb{R}^n$ and satisfies the inner sphere property of radius R. Suppose, moreover, that $H(\bar{x}(T), \nu) > 0$ for any $\nu \in N^P_{\mathscr{K}}(\bar{x}(T))$, $\nu \neq 0$, and T is differentiable at x_0. Then, we have that T is differentiable at $\bar{x}(t)$ for all $t \in [0, T(x_0))$.*

Proof. Showing that T is differentiable at a certain point z is equivalent to prove that $\partial^+ T(z)$ and $\partial^- T(z)$ are both nonempty. Thus, the conclusion in the case (b) comes from Theorem 4 together with [9, Theorem 4.1].

4.3 Second application: local C^2 regularity of the minimum time function

Theorem 2 and 3 apply to show that the existence of a proximal subgradient of $T(\cdot)$ at x is sufficient for the local regularity of $T(\cdot)$ in a neighborhood of the optimal trajectory starting form x. The proof is based upon ideas from [4, 7].

Theorem 6. *Assume $(SH), (A), (PC)$ and suppose that H is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$. Let $x_0 \in \mathscr{R} \setminus \mathscr{K}$ and let $\bar{x}$ be an optimal trajectory starting from x_0 reaching*

$\mathscr{K}$ at time $T(x_0)$. If $\partial^{-,P}T(x_0) \neq \emptyset$, then T is of class C^2 in a neighborhood of $\bar{x}([0,T(x_0)))$.

Proof. Recall first that the minimum time function T is semiconcave (see [10]). Thus, it is well known that since $\partial^{-,P}T(x_0) \neq \emptyset$, T must be differentiable at x_0. Therefore, the optimal trajectory for x_0 is unique, and we call it $\bar{x}$. It reaches the target at time $T(x_0)$. Set $\xi_0 := \bar{x}(T(x_0))$.
Thanks to assumption (PC), T is differentiable at $\bar{x}(t)$ for all $t \in [0,T(x_0))$ (see Theorem 5). Recall that and $-\nabla b_{\mathscr{K}}(\xi_0)$ is a proximal inner normal to $\mathscr{K}$ at ξ_0 with unit norm. Thus, $\overline{p}(t) = \nabla T(\bar{x}(t))$ where $\overline{p} : [0,T(x_0)] \to \mathbb{R}^n \setminus \{0\}$ is such that the pair $(\bar{x},\overline{p})$ solves: for a.e. $t \in [0,T(x_0)]$,

$$\begin{cases} -\dot{x}(t) = \nabla_p H(x(t),p(t)), \\ \dot{p}(t) \in \partial_x H(x(t),p(t)), \end{cases} \qquad p(T) = -\nabla b_{\mathscr{K}}(\xi_0) H(\xi_0, \nabla b_{\mathscr{K}}(\xi_0))^{-1}.$$

Set $Y(\xi_0,\cdot) = \bar{x}(T(x_0)-\cdot)$ and $P(\xi_0,\cdot) = \overline{p}(T(x_0)-\cdot)$. Thus, $(Y(\xi_0,\cdot),P(\xi_0,\cdot))$ solves (15).
Thanks to Theorem 2, it is sufficient to prove that the interval $[0,T(x_0)]$ does not contain any conjugate time for ξ_0. Let us proceed by contradiction, assuming that there exists a conjugate time $\bar{t}$ for ξ_0 with $\bar{t} \in (0,T(x_0))$. Fix $t \in (0,\bar{t})$. By Theorem 2, we deduce that there exists an open neighborhood $V_{\xi_0} \times U_t$ of (ξ_0,t) in $\partial\mathscr{K} \times \mathbb{R}$ such that the function $T(\cdot)$ is of class C^2 in the open neighborhood $Y(V_{\xi_0},U_t)$ of $Y(\xi_0,t)$. Furthermore, the Hessian of T is

$$D^2T(Y(\xi,s)) = R_{\xi,t}(\xi,s) := P_{\xi,t}(\xi,s)Y_{\xi,t}^{-1}(\xi,s),\ (\xi,s) \in V_{\xi_0} \times U_t, \tag{31}$$

where $(P_{\xi,t},Y_{\xi,t})$ is the solution to (16). Recall that equality (31) has to be understood in the sense that we have explained in Remark 4. Thus, for all $(\xi,s) \in V_{\xi_0} \times U_t$ we have that

$$\begin{aligned} &T(Y(\xi,s)) - T(Y(\xi_0,t)) - \langle \nabla T(Y(\xi_0,t)), Y(\xi,s) - Y(\xi_0,t)\rangle \\ &- \frac{1}{2}\langle R_{\xi,t}(\xi_0,t)(Y(\xi,s)-Y(\xi_0,t)), Y(\xi,s)-Y(\xi_0,t)\rangle = o(|Y(\xi,s)-Y(\xi_0,t)|^2). \end{aligned} \tag{32}$$

Moreover, since $-\overline{p}(0) \in \partial^{-,P}T(x_0)$, by Theorem 3 there exists $R_0 > 0$ and $c_0 \geq 0$ such that

$$T(y) - T(Y(\xi_0,t)) - \langle \nabla T(Y(\xi_0,t)), y - Y(\xi_0,t)\rangle \geq -c_0|y-Y(\xi_0,t)|^2, \tag{33}$$

whenever $y \in B(Y(\xi_0,t),R_0)$ and $t \in [0,T(x_0)]$. Without lost of generality, we can suppose that $Y(V_{\xi_0},U_t) \subset B(Y(\xi_0,t),R_0)$. Then, by (32) and (33),

$$\langle R_{\xi,t}(\xi_0,t)(Y(\xi,s)-Y(\xi_0,t)), Y(\xi,s)-Y(\xi_0,t)\rangle \geq -c_0|Y(\xi,s)-Y(\xi_0,t)|^2, \tag{34}$$

for all $(\xi,s) \in V_{\xi_0} \times U_t$. Since $Y(V_{\xi_0},U_t)$ is an open neighborhood of $Y(\xi_0,t)$ in $\mathbb{R}^n$, from (34) we deduce that

$$\langle R_{\xi,t}(\xi_0,t)\theta,\theta\rangle \geq -c_0, \quad \forall\, \theta \in S^{n-1}. \tag{35}$$

This provides a bound from below, uniform in $[0,\bar{t})$, of the quadratic form associated to $R_{\xi,t}(\xi_0,t)$. Furthermore, since T is semiconcave it holds that for any $\nu \in \mathbb{R}^n$ such that $\mid \nu \mid = 1$ we have $\frac{\partial^2 T}{\partial \nu^2} \leq C$ in the sense of distributions, where C is the semiconcavity constant of T (see, for instance, [13, Proposition 1.1.3]). Since T is twice differentiable on $Y(V_{\xi_0},U_t)$ for all $t \in [0,\bar{t})$, the distributional Hessian coincides with the classical Hessian (31) on such sets. We conclude that $R_{\xi,t}(\xi_0,t)$ must be bounded from above on $[0,\bar{t})$ by C, still in the sense of quadratic form. On the other hand, the operator norm of $R_{\xi,t}(\xi_0,t)$ goes to infinity as $t \to \bar{t}$, since we have supposed that $\bar{t}$ is conjugate for ξ_0. These facts together give a contradiction. Summarizing, we have proved that the all interval $[0,T(x_0)]$ does not contain conjugate times for ξ_0, and we conclude by Theorem 2 that the minimum time function T is of class C^2 in a neighborhood of $\bar{x}([0,T(x_0)))$.

Example 1. Here are some examples of systems with a smooth Hamiltonian. We consider a dynamic $\dot{x} = f(x,u)$ given by the *control-affine system* with drift:

$$f(x,u) = h(x) + F(x)u,\ u \in U, \tag{36}$$

where $U \subset \mathbb{R}^m$ is the closed unit ball, $h : \mathbb{R}^n \to \mathbb{R}^n$ is the drift and $F : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ is a matrix-valued function defined by m vector fields $f_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1,...m$, as $F(\cdot) = [f_1(\cdot),...,f_m(\cdot)]$. Suppose hereafter that the vector fields h and f_i, $i = 1,...,m$, are of class C^2 with sublinear growth. The Hamiltonian associated to this system is

$$H(x,p) = \langle p, h(x)\rangle + \mid F(x)^* p \mid .$$

We give the following two examples:

- We assume that $F(x)$ is surjective for all $x \in \mathbb{R}$ (that is, $m \geq n$ and that the rank of $F(x)$ is equal to n for all $x \in \mathbb{R}^n$). Note that if $F(x)$ is surjective, then $F(x)^*$ is injective. So its kernel is reduced to the singleton set $\{0\}$. Thus, for all $x,p \in \mathbb{R}^n$, $p \neq 0$,
$$H_p(x,p) = h(x) + \frac{F(x)F^*(x)p}{|F^*(x)p|},$$
$$H_{pp}(x,p) = \frac{F(x)F^*(x)}{|F^*(x)p\mid} - \frac{F(x)F(x)^*p \otimes F(x)F^*(x)p}{|F^*(x)p|^3}.$$
Thus, F satisfies assumption (SH) and H is clearly of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$. Note that also hypothesis $(H2)$ is satisfied. Similarly, we can consider strictly convex sets U with sufficiently smooth boundary.
Let us remark that the special case where $m = n$, $h \equiv 0$ and $F(x)$ is invertible for all $x \in \mathbb{R}^n$ corresponds to a Riemannian type problem. We are not in the standard Riemannian situation whenever $F(x)$ is not invertible, and in particular when $m < n$. We deal with subriemannnian type problems in the example below.
- We consider the control system (36) without drift, i.e. $h \equiv 0$. Let us first recall the notion of singular optimal trajectory for such control system. Let us introduce

the pre-Hamiltonian

$$H_0(x,p,u) = \sum_{i=1}^{m} u_i \langle p_i, f_i(x)\rangle,$$

that is smooth on $\mathbb{R}^n \times \mathbb{R}^n \times U$ because of the smoothness of the fields f_i. A trajectory $x : [0,T] \to \mathbb{R}^n$ reaching the target at time T (and corresponding to the control $u(\cdot)$) is said to be *singular* if and only if there exists an absolutely continuous arc $p : [0,T] \to \mathbb{R}^n \setminus \{0\}$ such that $(x(\cdot), p(\cdot), u(\cdot))$ solves the system

$$\begin{cases} -\dot{x}(t) = \nabla_p H_0(x(t), p(t), u(t)), \\ \dot{p}(t) = \nabla_x H_0(x(t), p(t), u(t)), \end{cases} \quad t \in [0,T], \tag{37}$$

and $p(\cdot)$ is orthogonal to each vector $f_1(x(\cdot)), ..., f_m(x(\cdot))$ on $[0,T]$ (i.e. $p(t) \in \ker F(x(t))^*$ for all $t \in [0,T]$). The control u is said to be singular as well.
Let $\overline{x}(\cdot)$ be a nonsingular optimal trajectory and let $\overline{p}(\cdot)$ a dual arc solving (37) and such that $\overline{p}(s) \notin \ker F(\overline{x}(s))^*$ for a time s in the interval $[0,T]$. Thus, even if $F(\cdot)$ is not surjective, in a neighborhood of $(\overline{x}(\cdot), \overline{p}(\cdot))$ the Hamiltonian H satisfies assumption (H) and is of class C^2, and $(\overline{x}(\cdot), \overline{p}(\cdot))$ solves also the system

$$\begin{cases} -\dot{x}(t) = \nabla_p H(x(t), p(t)), \\ \dot{p}(t) = \nabla_x H(x(t), p(t)), \end{cases} \quad t \in [0,T]. \tag{38}$$

To conclude, we recall that a sufficient condition to exclude the presence of non-trivial singular control is that the distribution corresponding to $f_1, ..., f_m$ is *fat*, i.e., the control system is *strongly bracket generating*. There is a vast literature devoted to this topic. The interested reader is referred to, e.g., [19, 20] for more details on geometric control.

Acknowledgement

This research is partially supported by the European Commission (FP7-PEOPLE-2010-ITN, Grant Agreement no. 264735-SADCO), and by the INdAM National Group GNAMPA. This work was completed while the first author was visiting the Institut Henri Poincaré and Institut des Hautes Études Scientifiques on a senior CARMIN position. The authors are grateful to the anonymous referee for her/his useful comments.

References

1. Aubin, J. P., Frankowska, H.: Set-valued analysis, Birkhäuser Boston Inc., Boston, MA (1990).
2. Cannarsa, P., Frankowska, H.: Some characterizations of optimal trajectories in control theory, *SIAM J. Control Optim.*, 29 (1991), pp. 1322-1347.

3. Cannarsa, P., Frankowska, H.: Interior sphere property of attainable sets and time optimal control problems, *ESAIM Control Optim. Calc. Var.*, 12 (2006), pp. 1322-1347.
4. Cannarsa, P., Frankowska, H.: From pointwise to local regularity for solutions of Hamilton-Jacobi-Bellman equations, *Calc. Var. Partial Differential Equations* 49 (2014), no. 3-4, 1061-1074.
5. Cannarsa, P., Frankowska, H.: Local regularity of the value function in optimal control, Systems & Control Letters, 62 (2013), pp. 791-794.
6. Cannarsa, P., Frankowska, H., Scarinci, T.: Sensitivity relations for the Mayer problem with differential inclusions, to appear in *ESAIM Control Optim. Calc. Var.*
7. Cannarsa, P., Frankowska, H., Scarinci, T.: Second-order sensitivity relations and regularity of the value funcion for Mayer's problem in optimal control (arXiv:1408.5354 [math.FA]).
8. Cannarsa, P., Frankowska, H., Sinestrari, C.: Optimality conditions and synthesis for the minimum time problem, *Set-Valued Anal.*, 8 (2000) pp. 127-148.
9. Cannarsa, P., Marigonda, A., Nguyen, K.T.: Optimality conditions and regularity results for the time optimal control problems with differential inclusions (arXiv:1311.4415 [math.OC]).
10. Cannarsa, P., Marino, F., Wolenski, P.: Semiconcavity of the minimum time function for differential inclusions, *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms*, 19 (2012), pp. 187-206.
11. Cannarsa, P., Marino, F., Wolenski, P.: The dual arc inclusion with differential inclusions, *Nonlinear Anal.*, 79 (2013), pp. 176-189.
12. Cannarsa, P., Nguyen, K. T.: Exterior Sphere Condition and Time Optimal Control for Differential Inclusions, *SIAM J. Control Optim.*, 49 (2011), pp. 2558-2576.
13. Cannarsa, P., Sinestrari, C.: Semiconcave functions, Hamilton-Jacobi equations, and optimal control, Birkhäuser Boston Inc. (2004).
14. Cannarsa, P., Wolenski, P.: Semiconcavity of the value function for a class of differential inclusions, *Discrete Contin. Dyn. Syst.*, 29 (2011), pp. 453-466.
15. Caroff N., Frankowska, H.: Conjugate points and shocks in nonlinear optimal control, *Trans. Amer. Math. Soc.* 348 (1996), pp. 3133-3153.
16. Clarke, F. H., Optimization and nonsmooth analysis, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA *SIAM*, Vol. 5. (1990).
17. Clarke, F., Vinter, R.: The relationship between the maximum principle and dynamic programming, *SIAM J. Control Optim.* 25 (1987) 1291-1311.
18. Frankowska, H. , Nguyen, L. V.: Local Regularity of the Minimum Time Function, *J. Optim. Theory and Appl.* (2014), pp. 1-24.
19. Jean, F.: Control of Nonholonomic Systems: from Sub-Riemannian Geometry to Motion Planning, Springer International Publishing, SpringerBriefs in Mathematics (2014).
20. Montgomery, R.: A Tour of Subriemannian Geometries, Their Geodesics and Applications, Mathematical Surveys and Monographs (2002).
21. Pignotti, C.: Rectifiability results for singular and conjugate points of optimal exit time problems, *Journal of Mathematical Analysis and Applications*, 270 (2002), n. 2, pp. 681-708.
22. Subbotina, N. N.: The maximum principle and the superdifferential of the value function, *Problems Control Inform. Theory/Problemy Upravlen. Teor. Inform.* 18 (1989) 151-160.
23. Vinter, R.: New results on the relationship between dynamic programming
24. Vinter, R.: Optimal control, *Systems & Control: Foundations & Applications*, Birkhäuser Boston, Inc., Boston, MA (2000).
25. Wolenski, P., Zhuang, Y.: Proximal analysis and the minimal time function, *SIAM J. Control Optim.* (1998) 36, n. 3, pp. 1048- 1072.