ON THE MINIMUM TIME FUNCTION AROUND THE ORIGIN

GIOVANNI COLOMBO AND KHAI T. NGUYEN

Abstract. We deal with finite dimensional linear and nonlinear control systems. If the system is linear and autonomous and satisfies the classical *normality* assumption, we improve the well known result on the strict convexity of the reachable set from the origin by giving a polynomial estimate. The result is based on a careful analysis of the switching function. We extend this result to nonautonomous linear systems, provided the time dependent system is not too far from the autonomous system obtained by taking the time to be 0 in the dynamics.
Using a linearization approach, we prove a bang-bang principle, valid in dimensions 2 and 3 for a class of nonlinear systems, affine and symmetric with respect to the control. Moreover we show that, for two dimensional systems, the reachable set from the origin satisfies the same polynomial strict convexity property as for the linearized dynamics, provided the nonlinearity is small enough. Finally, under the same assumptions we show that the epigraph of the minimum time function has positive reach, hence proving the first result of this type in a nonlinear setting. In all the above results, we require that the linearization at the origin be normal.
We provide examples showing the sharpness of our assumptions.

1. Introduction

In the theory of autonomous linear control systems, the assumption of *normality*, i.e., a strong controllability assumption requiring that if each control component is used separately, then Kalman rank condition is satisfied, is well known. In particular (see [15, Sections 14, 15, and 16]), normality implies that the control steering the origin to a point x in minimum time is unique and bang-bang; moreover the reachable set from the origin, $\mathcal{R}^\tau$, is a *strictly* convex body for all times $\tau > 0$. Simple examples, on the other hand, show that convexity of the reachable set is easily lost when passing to a nonlinear dynamics, even if the control covers all directions and appears linearly. In particular, the control system

$$\begin{cases} \dot y_1(t) &= -[y_2(t)]^2 + u_1(t), \quad -1 \le u_1 \le 1 \\ \dot y_2(t) &= u_2(t) \qquad\qquad\qquad\quad -1 \le u_2 \le 1 \end{cases} \tag{1.1}$$

which was analyzed in [6], fails to have convex (and even normally regular, i.e., "without inward corners") reachable sets from the origin at any positive time.
Similarly, the minimum time to reach the origin for a linear system is semiconvex, i.e., it is a quadratic perturbation of a convex function, provided a first order controllability assumption is satisfied (see [6, 7]), or has epigraph with positive reach, i.e., it satisfies a quite good kind of generalized convexity, provided Kalman rank condition holds (see [11, 12]). On the other hand, the same example introduced above shows that such properties may fail even for a simple nonlinear

Date: June 27, 2011.

2000 *Mathematics Subject Classification.* AMS 2000 MSC: 49N60, 49J52, 49E30.

Key words and phrases. Reachable sets, exterior sphere condition, Maximum Principle.

This work was partially supported by M.I.U.R., project "Viscosity, metric, and control theoretic methods for nonlinear partial differential equations" and by CARIPARO.

dynamics.

We are not aware of any result of semiconvexity type valid for a nonlinear dynamics, where actually results of semiconcave type are more natural and easier to obtain, provided the target (or the dynamics) satisfy an inner ball condition (see [7, 3, 13, 17, 4, 5]). Further results on the regularity of the minimum time function T, for two dimensional systems with a single input control, are described in [2, Chapter 3], where in particular, under generic assumptions, a characterization of smooth and nonsmooth points of the level sets of T is given.

In this paper, we give a contribution to the understanding of the behavior of the minimum time function both for linear, autonomous and nonautonomous, and nonlinear control systems. We take the origin as the target (or as the source, for the reversed dynamics), hence going outside the realm of semiconcavity, and give results, under easily verifiable assumptions, on the following topics:

- strict convexity of the reachable set from the origin
- uniqueness of the optimal control
- a nonlinear bang-bang principle
- extending backwards optimal trajectories
- positive reach of the epigraph of T.

Our method is based on new results on linear control systems satisfying the normality condition, and on linearization at the origin. The underlying idea, in fact, is requiring enough strength to such linearization, so that examples like (1.1) are ruled out. Our linear results hold in any space dimension, while the nonlinear part is confined to two or three dimensional spaces. Our arguments are based essentially on a careful analysis of the *switching function*, namely the function whose sign is expected to determine the optimal control, according to Pontryagin's Maximum Principle. The point is exactly showing that this sign is well defined, except for at most finitely many zeros. To this aim, the normality assumption is pivotal, as it permits to split any finite interval into finitely many sets, each one being a disjoint union of finitely many intervals, where the switching function or its derivatives are uniformly bounded away from zero (see Lemma 3.2). From this fact we are able to deduce a quantitative estimate on the strict convexity of the reachable set. More precisely (see Theorem 3.3), for a linear control system in $\mathbb{R}^N$ we show that for all $\tau > 0$ there exists a positive constant $\gamma > 0$ such that

$$\langle \zeta, y - x \rangle \leq -\gamma \|\zeta\| \, \|y - x\|^N \qquad \text{for all } x, y \in \mathcal{R}^\tau,\ \zeta \in N_{\mathcal{R}^\tau}(x)$$

(here $N_{\mathcal{R}^\tau}(x)$ denotes the normal cone to $\mathcal{R}^\tau$ at x). We show through an example that the exponent N is optimal. Section 3 is devoted to the above topic, together with an auxiliary study for a linear nonautonomous dynamics.

The nonlinear part starts with a nonlinear bang-bang result (Section 4), valid up to dimension 3. We consider a nonlinear control system which is affine with respect to the control: if the linearization at the origin is normal, then every optimal control is bang-bang. The proof is based on Pontryagin's Maximum Principle: if the nonlinearity contains only parts which are of order larger or equal to the space dimension, then we are able to transfer to the switching function all the properties satisfied by the switching function of the dynamics linearized at the origin. This idea is at the basis also of the strict convexity of the reachable set for a nonlinear two dimensional dynamics (see Theorem 5.1) and of proving that all points close enough to the origin are *optimal*, i.e., any trajectory steering a point to the origin optimally can be extended backwards still remaining optimal (see Theorem 6.2). In this case, the difficulty is extending the optimal control: our analysis permits to predict

backwards the sign of the switching function. Finally, we show under the same assumptions that the epigraph of the minimum time function T has positive reach, hence obtaining a rich bunch of regularity properties for T, as listed in Theorem 2.4. We also show through an example (see Example 5.3) that the assumptions on the nonlinear part cannot be avoided, while (1.1) shows that normality at the origin is essential. The restrictions on the space dimension for the nonlinear results will be explained after the relevant proofs (see Remarks 4.2 and 5.4).
To our knowledge, the results we present here as well as most of the used methods do not trace back to previous literature. In particular, the nonlinear bang-bang results of Krener [16] and Sussmann [19] seem to be of a very different nature.

2. Preliminaries

2.1. **Nonsmooth analysis and sets with positive reach.**

In all the paper, the space dimension N will be supposed larger or equal to 2.
Let $K \subset \mathbb{R}^N$ be closed. The distance function to K and the projection mapping onto K are defined respectively by

$$\begin{aligned} d_K(x) &= \min\{\|y-x\| \,:\, y \in K\}, \\ \pi_K(x) &= \{y \in K \,:\, \|y-x\| = d_K(x)\}. \end{aligned}$$

The boundary of K is $\operatorname{bdry} K$. Given $x \in K$ and $v \in \mathbb{R}^N$, we say that v is a *proximal normal* to K at x (and we will denote this fact by $v \in N_K(x)$) if there exists $\sigma = \sigma(v,x) \geq 0$ such that

$$\langle v, y-x\rangle \leq \sigma \|y-x\|^2 \qquad \text{for all } y \in K;$$

equivalently $v \in N_K(x)$ if and only if there exists $\rho > 0$ such that $\pi_K(x+\rho v) = \{x\}$, and in this case we say that the proximal normal v *is realized by a ball of radius* $\rho > 0$. If K is convex then $N_K(x)$ coincides with the normal cone of Convex Analysis. The proximal subdifferential of a function $f : \Omega \to \mathbb{R}$ at a point x of its domain, $\partial f(x)$, is the set of vectors $v \in \mathbb{R}^N$ such that

$$(v,-1) \in N_{\operatorname{epi}(f)}(x, f(x)),$$

where $\operatorname{epi}(f)$, the epigraph of f, is defined as

$$\operatorname{epi}(f) = \{(x,y) \in \Omega \times \mathbb{R} : y \geq f(x)\}.$$

For an introduction to nonsmooth analysis we make reference, e.g., to [10, Chapters 1 and 2].
We will make use of the following concepts due to Federer [14]: given an arbitrary set $K \subset \mathbb{R}^N$, we set

$$\begin{aligned} \operatorname{Unp}(K) &= \{x \in \mathbb{R}^N \,:\, \pi_K(x) \text{ is a singleton}\}, \\ \operatorname{reach}(K,x) &= \sup\{r \geq 0 \,:\, x + r\mathbb{B}^n \subseteq \operatorname{Unp}(K)\}, \end{aligned}$$

the latter being defined for $x \in K$. We remark that $\operatorname{reach}(K,x)$ is continuous with respect to $x \in K$.

Definition 2.1. *We say that a closed $K \subseteq \mathbb{R}^N$ has* positive reach *if*

$$\operatorname{reach}(K) := \inf_{x \in K} \operatorname{reach}(K,x) > 0.$$

A locally closed set K has locally positive reach *if* $\operatorname{reach}(K,x) > 0$ *for every $x \in K$.*

The positive reach property is actually an external sphere condition with (locally) uniform radius. More precisely, it holds (see [14, §4]):

Proposition 2.2. *Let $K \subset \mathbb{R}^N$ be closed. Then K has positive reach if and only if there exists a continuous function $\varphi : K \to [0, +\infty)$ such that the inequality*

$$\langle v, y - x\rangle \leq \varphi(x)\|v\| \, \|y - x\|^2 \tag{2.1}$$

holds for all $x, y \in K$ and $v \in N_K(x)$.

We say that $f : \Omega \to \mathbb{R}$ is of class $\mathcal{C}^{1,1}(\Omega)$ (here $\Omega \subseteq \mathbb{R}^N$ is open) if its partial derivatives exist and are Lipschitz.

Remark 2.3. *The case when $\varphi(x) \equiv 0$ in (2.1) is equivalent to the convexity of K, and (2.1) is in this sense a generalization of convexity. Furthermore, it is easy to see that if the boundary of K is the graph of a $\mathcal{C}^{1,1}$-function, then K has positive reach and φ in (2.1) can be taken as constant. Hence positive reach property generalizes $\mathcal{C}^{1,1}$-manifolds as well.*

Lower semicontinuous functions whose epigraph has positive reach enjoy remarkable regularity properties, which are similar to properties of convex functions. In particular, the following result holds true (we state it for continuous functions for simplicity). The Lebesgue N-dimensional measure and the Hausdorff d-dimensional measure are denoted, respectively, by $\mathcal{L}^N$ and $\mathcal{H}^d$.

Theorem 2.4. *Let $\Omega \subset \mathbb{R}^N$ be open, and let $f : \Omega \to \mathbb{R}$ be continuous, and such that epi(f) has locally positive reach. Then there exists a sequence of sets $\Omega_h \subseteq \Omega$ such that Ω_h is compact in Ω and*

(1) the union of Ω_h covers $\mathcal{L}^N$-almost all Ω;
(2) for all $x \in \bigcup_h \Omega_h$ there exist $\delta = \delta(x) > 0$, $L = L(x) > 0$ such that

$$f \text{ is Lipschitz on } B(x, \delta) \text{ with ratio } L, \text{ and hence semiconvex on } B(x, \delta).$$

Consequently,

(3) f is a.e. Fréchet differentiable and admits a second order Taylor expansion around a.e. point of its domain.

Moreover, the set of points where the graph of f is nonsmooth has small Hausdorff dimension. More precisely,

(4) for every $k = 1, \dots, N$, the set $\{x \in \operatorname{int}\operatorname{dom}(f) \mid$ the dimension of $\partial f(x)$ is $\geq k\}$ is countably $\mathcal{H}^{N-k}$-rectifiable.

Finally,

(5) f has locally bounded variation in Ω.

This result is essentially Theorem 5.1 and Proposition 7.1 in [11]. For properties of semiconvex/semiconcave functions we refer to [7].

2.2. **Control theory.** We will consider control systems linear or nonlinear with respect to the space variable and affine and symmetric with respect to the control. More precisely, we will consider the linear control system

$$\begin{cases} \dot{y}(t) \;=\; Ay(t) + Bu(t) & \text{a.e.} \\ u(t) \;\in\; \mathcal{U} = [-1, 1]^M & \text{a.e.} \\ y(0) \;=\; 0, \end{cases} \tag{2.2}$$

where $1 \le M \le N$ and $A \in \mathbb{M}_{N\times N}$, $B \in \mathbb{M}_{N\times M}$, being possibly time dependent, and $\mathcal{U} = [-1,1]^M \ni (u_1,\dots,u_M) =: u$, together with the nonlinear control system

$$\begin{cases} \dot y(t) \;=\; F(y(t)) + G(y(t))u(t) & \text{a.e.} \\ u(t) \;\in\; [-1,1]^M & \text{a.e.} \\ y(0) \;=\; 0, \end{cases} \tag{2.3}$$

where F and G are suitable vector fields (the actual assumptions will be stated later). We will use also the notation $B = (b_1,\dots,b_M)$ or $G = (g_1,\dots,g_M)$, where each entry is an N-dimensional column. We denote by $\mathcal{U}_{ad}$, the set of admissible controls, i.e., all measurable functions u, such that $u(s) \in \mathcal{U}$ for a.e. s. For any $u(\cdot) \in \mathcal{U}_{ad}$, the (unique, as it will follow from the assumptions on F and G) Carathéodory solution of (2.2) or of (2.3) is denoted by $y^u(\cdot)$. In the linear case,

$$y^u(t) = \int_0^t e^{A(t-s)}Bu(s)ds,$$

so that the reachable set from 0 in time t can be described by

$$\mathcal{R}^t = \Big\{ \int_0^t e^{A(t-s)}Bu(s)ds) \mid u(\cdot) \in \mathcal{U}_{ad} \Big\}. \tag{2.4}$$

It is well known that in the linear case the set $\mathcal{R}^t$ is convex and compact (see, e.g., [15, Lemma 12.1]), while in the nonlinear case (2.3) $\mathcal{R}^t := \{y^u(t) \mid u(\cdot) \in \mathcal{U}_{ad}\}$ is compact and not necessarily convex (see, e.g., [8, Chapter 10]).

For a fixed $x \in \mathbb{R}^N$, we define

$$\theta(x,u) := \min\,\{t \ge 0 \mid y^{x,u}(t) = 0\},$$

where $y^{x,u}(\cdot)$ denotes the solution of $\dot y = F(y) + G(y)u$ such that $y(0) = x$. Of course, $\theta(x,u) \in [0,+\infty]$, and $\theta(x,u)$ is the time taken for the trajectory $y^{x,u}(\cdot)$ to reach 0, provided $\theta(x,u) < +\infty$. The *minimum time* $T(x)$ to reach 0 from x is defined by

$$T(x) := \inf\,\{\theta(x,u) \mid u(\cdot) \in \mathcal{U}_{\rm ad}\}. \tag{2.5}$$

Observe that the sublevel $\mathcal{R}_t = \{x : T(x) \le t\}$ of $T(\cdot)$ equals the reachable set from the origin within the (same) time t for the *reversed dynamics*

$$\dot x = -F(x) - G(x)u, \quad u \in \mathcal{U}.$$

If $\bar u$ is an admissible control steering x to the origin in the minimum time $T(x)$ (i.e., an *optimal* control), then the Dynamic Programming Principle (see, e.g., Proposition 2.1, Chapter IV, in [1]) implies that $T(\cdot)$ is strictly increasing along the optimal trajectory $y^{x,\bar u}$. Therefore, for all $0 < t < T(x)$ the point $y^{x,\bar u}(t)$ belongs to the boundary of $\mathcal{R}_t$.

Pontryagin's Maximum Principle is a fundamental tool for the analysis of optimal control problems. We state it for points belonging to the boundary of reachable sets. In view of the previous remark, this will apply also to points belonging to optimal trajectories. We give first its linear version.

Theorem 2.5 (Pontryagin's Maximum Principle for linear systems)**.** *Consider the problem* (2.2)*, fix $T > 0$, and suppose $\bar x \in \mathcal{R}^T$ is realized by the control $\bar u(\cdot) \in \mathcal{U}_{ad}$ (i.e., $y^{\bar u}(T) = \bar x$). Then $\bar x \in \operatorname{bdry}\mathcal{R}^T$ if and only if for some $\zeta \in N_{\mathcal{R}^T}(\bar x)$, $\zeta \ne 0$, it holds*

$$\bar u_i(t) = \operatorname{sign}\langle \zeta, e^{A(T-t)}b_i\rangle, \quad \text{a.e. } t \in [0,T], \tag{2.6}$$

for all $i = 1,2,\dots,M$.

A well known reference for this result is [15, Lemma 13.1].

Before stating Pontryagin's principle for the nonlinear case (2.3) (with $M = 1$), we need to introduce the *Maximized Hamiltonian*.
We define for every triple $(x, p, u) \in \mathbb{R}^N \times \mathbb{R}^N \times [-1, 1]$

$$\mathcal{H}(x, p, u) = \langle p, F(x)\rangle + u\langle p, G(x)\rangle$$

and

$$H(x, p) = \max\{\mathcal{H}(x, p, u) : u \in [-1, 1]\}.$$

Then Pontryagin's principle reads as follows (see, e.g., [2, Section 2.1] or [9, Theorem 3.5.4]).

Theorem 2.6 (Pontryagin's Maximum Principle for nonlinear systems)**.** *Fix $T > 0$ and let y^u be a trajectory of (2.3) such that $y^u(T)$ belongs to the boundary of $\mathcal{R}^T$. Then there exist a function $\lambda : [0, T] \to \mathbb{R}^N$, never vanishing, and a constant $\lambda_0 \le 0$ such that for a.e. $t \in [0, T]$ one has:*

i) $\dot\lambda(t) = -\lambda(t)\big(DF(y^u(t)) + DG(y^u(t))u(t)\big)$,
ii) $\mathcal{H}(y^u(t), \lambda(t), u(t)) + \lambda_0 = 0$,
iii) $\mathcal{H}(y^u(t), \lambda(t), u(t)) = H(y^u(t), \lambda(t))$.

Furthermore,

$$\lambda(T) \in N_{\mathcal{R}^T}(y^u(T))$$

where here $N_{\mathcal{R}^T}(y^u(T))$ denotes the Clarke normal cone.

Definition 2.7. *We say that a control u is essentially determined by Pontryagin's Principle if for any u_1 satisfying iii) in Theorem 2.6 (for the adjoint curve λ associated with the trajectory y^u) one has $u_1(t) = u(t)$ a.e. in $[0, T]$.*

In the following we will make extensive use of the classical concept of normality for linear systems, which we are now going to introduce.

Definition 2.8. *The system (2.2) is normal if and only if for every column b_i of B, $i = 1, \dots, M$, we have*

$$\operatorname{Rank}\big[b_i, Ab_i, \dots, A^{N-1}b_i\big] = N.$$

The main classical result for normal linear systems is concerned with the reachable set.

Theorem 2.9. *Assume that the linear control system(2.2) is normal. Then the reachable set $\mathcal{R}^T$ is strictly convex for any $T > 0$.*

Proof. One can find a proof in [15], Sections 14 and 15. □

Remark 2.10. *If the system (2.2) is normal then (A, B) satisfies the Kalman rank condition. Therefore the minimum time function is everywhere finite and continuous (actually Hölder continuous with exponent $1/N$, see, e.g., Theorem 17.3 in* [15] *and Theorem 1.9, Chapter IV, in* [1] *and references therein).*

3. Quantitative strict convexity of reachable sets: the linear case

3.1. **Autonomous systems.**

This subsection is devoted to improving the classical result on the strict convexity of reachable sets for normal linear control systems of the type (2.2). We will give an estimate for the boundary of reachable sets which implies a uniform (polynomial) strict convexity with an optimal exponent.
In the first Lemma we define the switching function and begin studying its behavior under the normality assumption.

Lemma 3.1. *Let $A \in \mathbb{M}_{N\times N}$ and $b \in \mathbb{R}^N$ be such that*

$$\text{Rank}\big[b, Ab, \dots, A^{N-1}b\big] = N. \tag{3.1}$$

Take $\zeta \in \mathbb{R}^N$, with $\|\zeta\| = 1$, and define, for $s \in [0, +\infty)$

$$g(s) \;=\; \langle e^{As}b \;,\; \zeta\rangle. \tag{3.2}$$

Then there exists a constant $\mathcal{L}$, depending only on A, b, N such that, for all $s \in [0, +\infty)$,

$$\sum_{i=0}^{N-1} |g^{(i)}(s)| \geq \mathcal{L}e^{-\|A\|s}. \tag{3.3}$$

Proof. Set

$$K = \Big(b,\;\; Ab,\;\; \dots,\;\; A^{N-1}b\Big)$$

and observe that, by (3.1)

$$\mathcal{L} = \min_{\|\zeta\|=1} \;\|K\zeta\| \;> 0. \tag{3.4}$$

Fix $\zeta \in \mathbb{R}^N$ with $\|\zeta\| = 1$ and write $\zeta_1(s) = e^{sA^T}\zeta$. Observe that $\zeta = e^{-A^T s}\zeta_1(s)$ and $\|\zeta_1(s)\| \geq e^{-s\|A\|}$. We compute now, for $i = 0, 1, \dots, N-1$,

$$g^{(i)}(s) = \langle e^{As}A^i b, \zeta\rangle = \langle A^i b, \zeta_1(s)\rangle. \tag{3.5}$$

Therefore,

$$\begin{aligned} K\zeta_1(s) \;&=\; \Big(b,\;\; Ab,\;\; \dots,\;\; A^{N-1}b\Big)\,\zeta_1(s) \\ &=\; \Big(g^{(0)}(s),\;\; g^{(1)}(s),\;\; \dots,\;\; g^{(N-1)}(s)\Big). \end{aligned}$$

Using (3.4) we have that

$$\|K\zeta_1(s)\| \geq \mathcal{L}e^{-s\|A\|}.$$

On the other hand,

$$\|K\zeta_1(s)\| \leq \sum_{i=0}^{N-1} |g^{(i)}(s)|$$

and the proof is concluded. □

The next Lemma is crucial for estimating the number of zeros of the switching function g (corresponding to the number of switching points of the optimal control associated with g) and for studying their multiplicity. We recall that the constant $\mathcal{L}$ was defined in (3.4).

Lemma 3.2. *Let $A \in \mathbb{M}_{N\times N}$ and $b \in \mathbb{R}^N$ be satisfying (3.1). Take $\zeta \in \mathbb{R}^N$, $\|\zeta\| = 1$, and fix $T > 0$. Let $g(s)$, $s \in [0,T]$, be defined as in (3.2).*

Then there exist disjoint sets $I_0, \dots, I_{N-1}$ and numbers $\mathcal{N}_i$, depending only on A, b, T and N such that

$$[0,T] = \bigcup_{i=0}^{N-1} I_i$$

and, for all $i = 0, 1, \dots, N-1$, the set I_i is the disjoint union of at most $\mathcal{N}_i$ intervals. Moreover, for each $i = 0, 1, \dots, N-1$, for all $s \in I_i$, we have

$$|g^{(i)}(s)| \geq \frac{\mathcal{L}}{N} e^{-\|A\|s}. \tag{3.6}$$

Proof. We proceed by induction for i from 0 to $N-1$. Set

$$c(s) = \frac{\mathcal{L} e^{-\|A\|s}}{N}, \tag{3.7}$$

and

$$J_0 = \{s \in (0,T) \mid |g(s)| < c(s)\}.$$

Since J_0 is open, we can write it as the disjoint union of at most countably many open intervals,

$$J_0 = \bigcup_{k=1}^{\infty} J_0^k. \tag{3.8}$$

We assume, without loss of generality, that there are at least N such intervals. Fix now any number $N' \geq N$, and take a subfamily of the intervals J_0^k consisting of at most N' elements. Without loss of generality, we can rearrange their indexes k so that $J_0^k = (a_{2k}, a_{2k+1})$, where $1 \leq k \leq N'$ and $0 \leq a_2 < a_3 \leq a_4 < a_5 \leq \ldots \leq a_{2k} < a_{2k+1} \ldots < a_{2N'+1} \leq T$.

Now, fix k and consider the N intervals $(a_{2k}, a_{2k+1}), \dots, (a_{2(k+N-1)}, a_{2(k+N)-1})$. Set, for $j = 0, 1, \dots, N-1$,

$$(a_{2(k+j)}, a_{2(k+j)+1}) := I_j^-,$$

and, for $j = 0, 1, \dots, N-2$

$$[a_{2(k+j)+1}, a_{2(k+j+1)}] := I_j^+.$$

Observe that

$$\bigcup_{j=0}^{N-1} I_j^- \cup \bigcup_{j=0}^{N-2} I_j^+ = (a_{2k}, a_{2(k+N)-1}).$$

We are going to give a lower bound on $|a_{2(k+N)-1} - a_{2k}|$ which will turn out to be independent of both k and N'. From this fact it will follow automatically that the intervals (a_{2k}, a_{2k+1}) are nonempty only for finitely many k.

Observe that for each $j = 0, 1, \dots, N-2$, there exists at least one point $c_j^1 \in I_j^+$ such that $g'(c_j^1) = 0$. Therefore, there exist at least $N-2$ points, say c_j^2, for $j = 0, 1, \dots, N-3$, such that

$$c_j^2 \in (c_j^1, c_{j+1}^1) \quad \text{and} \quad g''(c_j^2) = 0.$$

Proceeding by induction we see that, for each $i = 1, \dots, N-1$, there exists at least one point $c_i \in (a_{2k+1}, a_{2(k+N)-1})$ such that $g^{(i)}(c_i) = 0$.
Pick any $s_0 \in (a_{2k}, a_{2k+1})$. We have

$$|g(s_0)| < c(s_0), \tag{3.9}$$

and, for $i = 1, \dots, N-1$,

$$\begin{aligned} |g^{(i)}(s_0)| &= |g^{(i)}(s_0) - g^{(i)}(c_i)| = \Big| \int_{s_0}^{c_i} g^{(i+1)}(s) ds \Big| \\ &\leq \int_{a_{2k}}^{a_{2(k+N)-1}} |g^{(i+1)}(s)| ds \leq (a_{2(k+N)-1} - a_{2k}) e^{\|A\|T} \|A^{i+1} b\|, \end{aligned}$$

where the last inequality is due to (3.5). Therefore,

$$\sum_{i=1}^{N-1} |g^{(i)}(s_0)| \leq (a_{2(k+N)-1} - a_{2k}) e^{\|A\|T} \sum_{i=1}^{N-1} \|A^{i+1} b\|. \tag{3.10}$$

On the other hand, recalling (3.3), (3.7), and (3.9) we have

$$\sum_{i=1}^{N-1} |g^{(i)}(s_0)| \geq \mathcal{L} e^{-\|A\| s_0} - c(s_0) = \frac{N-1}{N} \mathcal{L} e^{-\|A\| s_0} \geq \frac{N-1}{N} \mathcal{L} e^{-\|A\|T}. \tag{3.11}$$

From (3.10) and (3.11) we obtain

$$a_{2(k+N)-1} - a_{2k} \geq \frac{(N-1)\mathcal{L} e^{-2\|A\|T}}{N \sum_{i=1}^{N-1} \|A^{i+1} b\|}, \tag{3.12}$$

which is the desired estimate. Observe that the right hand side of (3.12) depends only on A, b, T and N.

We set now $\mathcal{N}_0$ to be the number of nonempty intervals contributing to the union in (3.8), and recall that we have just proved that $\mathcal{N}_0$ depends only on A, b, T and N, and actually

$$\mathcal{N}_0 \leq \frac{N^2}{N-1} \frac{T}{\mathcal{L}} e^{2\|A\|T} \sum_{i=1}^{N-1} \|A^{i+1} b\| + N - 1. \tag{3.13}$$

Set $I_0 = [0, T] \backslash J_0$ and observe that we have completed the proof of the lemma for $i = 0$.

After this step, we formulate our induction process. We are going to construct, for each $i = 0, \dots, N-1$, two disjoint sets I_i, J_i with the following properties

(ind1) for every $s \in I_i$, $|g^{(i)}(s)| \geq c(s)$;
(ind2) for every $s \in J_i$, $\sum_{j=i+1}^{N-1} |g^{(j)}(s)| \geq (N-i-1) c(s)$;
(ind3) $J_i \bigcup I_i = J_{i-1}$;
(ind4) J_i is a finite union of open intervals whose number is at most $\mathcal{N}_i$, and $\mathcal{N}_i$ depends only on $T, \mathcal{L}, A, b, N, i$;
(ind5) I_i is the finite union of at most $\mathcal{N}_i + \mathcal{N}_{i-1}$ intervals.

For $i = 0$ the above construction was already performed (take $J_{-1} = (0, T)$). Pick any $i = 1, \dots, N-2$ (the case $i = N-1$ will be treated separately) and assume that (ind1), …, (ind5) hold up to $i-1$. We wish to show that the above statements hold for i as well. To this aim, consider the set

$$J_i := \{ s \in J_{i-1} \mid |g^{(i)}(s)| < c(s) \}.$$

For every connected component (a, b) of J_{i-1}, we are going to prove that $J_i^{(a,b)} := J_i \cap (a, b)$ is a finite union of intervals, and give a bound on their number $\mathcal{N}_i^{(a,b)}$.

So, fix a connected component (a, b) of J_i and represent the open set $J_i^{(a,b)}$ as a disjoint union of at most countably many intervals J_i^k, $k \in \mathbb{N}$ (for simplicity of writing we drop the dependence on (a, b)). Assume without loss of generality that there are at most $N - i$ intervals J_i^k, fix any number $N'' \geq N - i$, and take any subfamily of $\{J_i^k\}$ consisting of at most N'' intervals. We can write

$J_i^k = (a_{2k}, a_{2k+1})$, where $1 \le k \le N''$ and $a \le a_2 < a_3 \le a_4 < a_5 \le \ldots < a_{2N''+1} \le b$. Fix k and consider the intervals $(a_{2k}, a_{2k+1}), \ldots, (a_{2(k+N-i-1)}, a_{2(k+N-i)-1})$. Set, for $j = 0, \ldots, N-i-1$,

$$(a_{2(k+j)}, a_{2(k+j)+1}) := I_j^-,$$

and, for $j = 0, \ldots, N-i-2$,

$$[a_{2(k+j)+1}, a_{2(k+j+1)}] := I_j^+.$$

Observe that

$$\bigcup_{J=0}^{N-i-1} I_j^- \cup \bigcup_{j=1}^{N-i-2} I_j^+ = (a_{2k}, a_{2(k+N-i)-1}).$$

For each $j = 0, \ldots, N-i-2$ there exists at least one point $c_j^i \in I_j^+$ such that $g^{(i)}(c_j^i) = 0$. Proceeding by induction we see that for each $m = 0, \ldots, N-i-2$ there exists a point $c_m \in (a_{2k+1}, a_{2(k+N-i)-1})$ such that $g^{(i+m)}(c_m) = 0$. Pick any $s_0 \in (a_{2k}, a_{2(k+1)})$. By arguing as for $i = 0$, we obtain on one hand

$$|g^{(i)}(s_0)| < c(s_0) \tag{3.14}$$

and, for all $m = 0, \ldots N-i-1$,

$$|g^{(i+m)}(s_0)| \le (a_{2(k+N-i)-1} - a_{2k}) e^{\|A\|T} \left\|A^{i+m+1} b\right\|,$$

the latter inequality being due to (3.5). Thus

$$\sum_{m=i+1}^{N-1} |g^{(m)}(s_0)| \le (a_{2(k+N-i)-1} - a_{2k}) e^{\|A\|T} \sum_{m=i+1}^{N-1} \left\|A^{m+1} b\right\|. \tag{3.15}$$

On the other hand, owing to (ind2) and (3.14), we obtain

$$\sum_{m=i+1}^{N-1} |g^{(m)}(s)| \ge (N-i-1) c(s_0).$$

By combining the above inequality with (3.15) we now obtain

$$a_{2(k+N-i)-1} - a_{2k} \ge \frac{N-i-1}{N} \mathcal{L} e^{-\|A\|T} \frac{1}{\sum_{m=i+1}^{N-1} \|A^{m+1} b\|}.$$

Therefore, $J_i^{(a,b)}$ is the union of finitely many disjoint open intervals $\left(a_{2k}^{i+1}, a_{2k+1}^{i+1}\right)$, $k = 1, \ldots, \mathcal{N}_i^{(a,b)}$, where

$$\mathcal{N}_i^{(a,b)} \le \frac{N(N-i)}{N-i-1} \frac{|b-a|}{\mathcal{L}} e^{2\|A\|T} \sum_{m=i+1}^{N-1} \|A^{m+1} b\| + N - i - 1. \tag{3.16}$$

We define

$$I_i^{(a,b)} = \{s \in (a,b) \mid |g^{(i)(s)}| \ge c(s)\}$$

and observe that $I_i^{(a,b)}$ is the union of at most $\mathcal{N}_i^{(a,b)} + 1$ intervals.
We finally set I_i to be the union of the $I_i^{(a_j,b_j)}$ over all the (at most $\mathcal{N}_{i-1}$) connected components (a_j, b_j) of J_{i-1}. Therefore, I_i is the union of at most $\mathcal{N}_i + \mathcal{N}_{i-1}$ intervals, where

$$\mathcal{N}_i = \sum_{j=1}^{\mathcal{N}_{i-1}} \mathcal{N}_i^{(a_j,b_j)} \le \frac{N(N-i)}{N-i-1} \frac{T}{\mathcal{L}} e^{2\|A\|T} \sum_{m=i+1}^{N-1} \|A^{m+1} b\| + \mathcal{N}_{i-1}(N-i-1). \tag{3.17}$$

Finally we observe that J_i is the union of at most $\mathcal{N}_i$ open intervals.

If $i = N-1$, we observe that for each $s \in J_{N-2}$, recalling (ind2) we have $|g^{(N-1)}(s)| \geq c(s)$. Therefore we set $J_{N-1} = \emptyset$ and $I_{N-1} = J_{N-2}$. The proof is concluded. □

We are now going to prove the main result of this subsection.

Theorem 3.3. *Consider the linear control system*

$$\dot{x} = Ax + Bu, \tag{3.18}$$

where $1 \leq M \leq N$, $A \in \mathbb{M}_{N\times N}$, $B \in \mathbb{M}_{N\times M}$, and $u = (u_1, \dots, u_M) \in [-1,1]^M$.
Assume that (3.18) is normal, i.e., for every column b_j, $j = 1, \dots, M$, of B,

$$\text{rank } [b_j, Ab_j, \dots, A^{N-1}b_j] = N.$$

Let $\mathcal{R}^T$ be defined according to (2.4). Then for all $T > 0$ there exists a constant $\gamma > 0$, depending only on N, M, A, B, T such that for all $x, y \in \mathcal{R}^T$, for all $\zeta \in N_{\mathcal{R}^T}(x)$, the inequality

$$\langle \zeta, y - x\rangle \leq -\gamma \, \|\zeta\| \, \|y - x\|^N \tag{3.19}$$

holds. Moreover, there exists another constant γ', depending only on N, M, A, B, T, such that

$$\text{the ball } \; B(0, \gamma' T^N) \; \text{ is contained in } \mathcal{R}^T. \tag{3.20}$$

for all $T > 0$. Finally, the constants γ and γ' are bounded away from zero as $T \to 0^+$.

Proof. We consider first the case $M = 1$, so (3.18) reads as

$$\dot{x} = Ax + bu \quad , \; |u| \leq 1,$$

for a suitable $b \in \mathbb{R}^N$. Fix $\bar{x} \in \text{bdry}\, \mathcal{R}^T$ together with an optimal control $\bar{u}(\cdot)$ steering 0 to $\bar{x}$ in time T and $\zeta \in N_{\mathcal{R}^T}(\bar{x})$, $\|\zeta\| = 1$. We assume first that ζ satisfies Pontryagin's maximum principle, i.e., for a.e. $t \in [0,T]$,

$$\bar{u}(t) = \text{sign}\langle \zeta, e^{A(T-t)}b\rangle$$

and prove (3.19) for such ζ. The general case will then follow using Proposition 7.3.
So, we take $\bar{y} \in \mathcal{R}^T$ together with a control $u(\cdot)$ steering 0 to $\bar{y}$ and compute:

$$\begin{aligned} \langle \zeta, \bar{y} - \bar{x}\rangle &= \int_0^T \langle \zeta, e^{A(T-t)}b\rangle (u(t) - \bar{u}(t))\, dt \\ &\qquad \text{(recalling (2.6))} \\ &= -\int_0^T |\langle \zeta, e^{A(T-t)}b\rangle| \, |u(t) - \bar{u}(t)|\, dt. \end{aligned}$$

Set $K(t) = \frac{1}{2}|u(t) - \bar{u}(t)|$ and observe that $0 \leq K(t) \leq 1$ for a.e. $t \in [0,T]$, and

$$\langle \zeta, \bar{y} - \bar{x}\rangle = -2\int_0^T |\langle \zeta, e^{A(T-t)}b\rangle| K(t) dt = -2\int_0^T |\langle \zeta, e^{At}b\rangle| K(T-t) dt. \tag{3.21}$$

Moreover,

$$\|\bar{y} - \bar{x}\| = \left\| \int_0^T e^{A(T-t)}b(u(t) - \bar{u}(t)) dt \right\| \leq 2e^{T\|A\|}\|b\| \int_0^T K(t) dt. \tag{3.22}$$

Set, for $s \in [0, +\infty)$,

$$g(s) = \langle e^{As}b \; , \; \zeta\rangle.$$

By Lemma 3.2 there exist disjoint sets $I_0, I_1, \dots, I_{N-1}$ and numbers $\mathcal{N}_i$ such that $[0,T] = \bigcup_{i=0}^{N-1} I_i$, each I_i is the disjoint union of at most $\mathcal{N}_i$ intervals and (3.6) holds. Observe that, in particular, it follows that g may vanish at most at finitely many times, and so, recalling (2.6), the control $\bar{u}$ is piecewise constant and equal to either 1 or to -1.
We rewrite

$$\langle \zeta, \bar{y} - \bar{x} \rangle = -2 \sum_{i=0}^{N-1} \int_{I_i} |g(s)| K_1(s) ds \tag{3.23}$$

where $K_1(s) = K(T-s)$. We are now going to estimate separately the integrals $\int_{I_i} |g(s)| K_1(s) ds$, for all $i = 0, 1, \dots, N-1$.
For $i = 0$, we have

$$\int_{I_0} |g(s)| K_1(s) ds \geq \frac{\mathcal{L}}{N} e^{-\|A\|T} \int_{I_0} K_1(s) ds. \tag{3.24}$$

Fix now $i = 1, 2, \dots, N-1$, and write, recalling Lemma 3.2,

$$\overline{I_i} = \bigcup_{j=1}^{\mathcal{N}_i} [a_{ij}, b_{ij}]$$

where all *open* intervals (a_{ij}, b_{ij}) are disjoint. Recalling (3.6), we have, for all $s \in I_i$, $|g^{(i)}(s)| \geq \frac{\mathcal{L}}{N} e^{-\|A\|T}$. Fix $j \in \{1, 2, \dots, \mathcal{N}_i\}$. We are now going to apply inductively Lemma 7.2 on $[a_{ij}, b_{ij}]$ with the functions $g^{(i-k-1)}$ in place of f, for $k = 0, \dots, i-1$. Let $k = 0$ and set $f = g^{(i-1)}$. Then the assumption (7.1) is satisfied with $C = \frac{\mathcal{L}}{N} e^{-\|A\|T}$, thanks to (3.6), and Lemma 7.2 yields that for some point $c_{ij}^0 \in [a_{ij}, b_{ij}]$ we have

$$|g^{(i-1)}(s)| \geq C |s - c_{ij}^0| \quad \forall s \in [a_{ij}, b_{ij}].$$

Let $k = 1$. By applying Lemma 7.2 on each of the two (possibly degenerate) intervals a_{ij}, c_{ij}^0, c_{ij}^0, b_{ij} to the function $f = g^{(i-2)}$ with $C = \frac{\mathcal{L}}{N} e^{-\|A\|T}$, we find suitable points $c_{ij}^1 \in [a_{ij}, c_{ij}^0]$ and $c_{ij}^2 \in [c_{ij}^0, b_{ij}]$ such that we have both

$$|g^{(i-2)}(s)| \;\geq\; \frac{C}{2}(s - c_{ij}^1)^2 \quad \forall s \in [a_{ij}, c_{ij}^0]$$

and

$$|g^{(i-2)}(s)| \;\geq\; \frac{C}{2}(s - c_{ij}^2)^2 \quad \forall s \in [c_{ij}^0, b_{ij}].$$

By continuing the induction process until $k = i-1$, we split the interval $[a_{ij}, b_{ij}]$ into at most 2^i intervals $[a_{ij} = c_{ij}^0, c_{ij}^1], [c_{ij}^1, c_{ij}^2], \dots, [c_{ij}^{2^i-1}, b_{ij} = c_{ij}^{2^i}]$ (some of them being possibly degenerate) such that for all $l = 0, 1, \dots, 2^i - 1$ and $s \in [c_{ij}^l, c_{ij}^{l+1}]$ one has

$$\text{either} \;\; |g(s)| \geq \frac{C}{i!}(s - c_{ij}^l)^i \;\; \text{or} \;\; |g(s)| \geq \frac{C}{i!}(c_{ij}^{l+1} - s)^i. \tag{3.25}$$

Recalling (3.23) and the above discussion, we have

$$\begin{aligned} \langle \zeta, \bar{y} - \bar{x} \rangle &= -2 \sum_{i=0}^{N-1} \int_{I_i} |g(s)| K_1(s)\, ds \\ &= -2 \left[\int_{I_0} |g(s)| K_1(s)\, ds + \sum_{i=1}^{N-1} \sum_{j=0}^{\mathcal{N}_i - 1} \sum_{l=0}^{2^i - 1} \int_{c_{ij}^l}^{c_{ij}^{l+1}} |g(s)| K_1(s)\, ds \right]. \end{aligned}$$

Recalling (3.24) and (3.25), and $C = \frac{\mathcal{L}}{N} e^{-\|A\|T}$, we obtain from the above expression that

$$\langle \zeta, \bar{y} - \bar{x} \rangle \leq -\frac{2\mathcal{L}}{N} e^{-\|A\|T} \left[\int_{I_0} K_1(s)\, ds + \sum_{i=1}^{N-1} \sum_{j=0}^{\mathcal{N}_i - 1} \sum_{l=0}^{2^i-1} \int_{c_{ij}^l}^{c_{ij}^{l+1}} \frac{|s - \bar{c}_{ij}^l|^i}{i!} K_1(s)\, ds \right] \tag{3.26}$$

where $\bar{c}_{ij}^l$ is either c_{ij}^l or c_{ij}^{l+1}, according to the two possibilities appearing in (3.25). Applying Lemma 7.1 to each summand of (3.26) we therefore obtain

$$\begin{aligned}
\langle \zeta, \bar{y} - \bar{x} \rangle &\leq -2\frac{\mathcal{L}}{N} e^{-\|A\|T} \left[\int_{I_0} K_1(s)\, ds + \sum_{i=1}^{N-1} \sum_{j=0}^{\mathcal{N}_i - 1} \sum_{l=0}^{2^i-1} \frac{\big(\int_{c_{ij}^l}^{c_{ij}^{l+1}} K_1(s)\, ds \big)^{i+1}}{(i+1)!} \right] \\
&\qquad \text{(using the convexity of } x \mapsto x^{i+1} \text{ on the positive half line)} \\
&\leq -\frac{2\mathcal{L}}{N} e^{-\|A\|T} \left[\int_{I_0} K_1(s)\, ds + \sum_{i=1}^{N-1} \sum_{j=0}^{\mathcal{N}_i - 1} \frac{1}{(i+1)!\ 2^{i^2}} \Big(\int_{a_{ij}}^{b_{ij}} K_1(s)\, ds \Big)^{i+1} \right] \\
&\leq -\frac{2\mathcal{L}}{N} e^{-\|A\|T} \left[\int_{I_0} K_1(s)\, ds + \sum_{i=1}^{N-1} \frac{1}{(i+1)!\ 2^{i^2} \mathcal{N}_i^i} \Big(\int_{I_i} K_1(s)\, ds \Big)^{i+1} \right].
\end{aligned}$$

Thus, recalling that $0 \leq K_1(s) \leq 1$ a.e.,

$$\begin{aligned}
\langle \zeta, \bar{y} - \bar{x} \rangle \ \leq\ -\frac{2\mathcal{L}}{N} e^{-\|A\|T} \Bigg[& \frac{1}{|I_0|^{N-1}} \Big(\int_{I_0} K_1(s)\, ds \Big)^N \\
& + \sum_{i=1}^{N-1} \frac{1}{(i+1)!\ 2^{i^2} \mathcal{N}_i^i |I_i|^{N-i-1}} \Big(\int_{I_i} K_1(s)\, ds \Big)^N \Bigg].
\end{aligned} \tag{3.27}$$

Owing to (3.13) and (3.17), we see that $\mathcal{N}_0 < \mathcal{N}_1 < \ldots < \mathcal{N}_N \leq C(A, b, N) e^{2\|A\|T}$ where $C(A, b, N)$ depends only on A, b, N. Therefore, we obtain finally from (3.27) and the definition of K_1 that

$$\langle \zeta, \bar{y} - \bar{x} \rangle \leq -C(A, b, N, T) e^{-\|A\|T} \Big(\int_0^T |u(t) - \bar{u}(t)| \Big)^N, \tag{3.28}$$

where $C(A, b, N, T)$ is a positive constant, depending only on A, b, N, T such that

$$\liminf_{T \to 0} C(A, b, N, T) > 0. \tag{3.29}$$

Recalling (3.22), we complete the proof for the case $M = 1$ (i.e., a scalar control) by setting

$$\gamma = 2^{-N} e^{-\|A\|(N+1)T} \|b\|^{-N} C(A, b, N, T).$$

Let now $M > 1$. Take $\bar{x} \in \operatorname{bdry} \mathcal{R}^T$ together with an optimal control $\bar{u}(\cdot) = (\bar{u}_1(\cdot), \ldots, \bar{u}_M(\cdot))$ steering the origin to $\bar{x}$ in the optimal time T, and $\bar{y} \in \mathcal{R}^T$ with a control $u(\cdot) = (u_1(\cdot), \ldots, u_M(\cdot))$ steering the origin to $\bar{y}$ in time T. Then, for each $\zeta \in N_{\mathcal{R}^T}(\bar{x})$, $\|\zeta\| = 1$, we can write

$$\langle \zeta, \bar{y} - \bar{x} \rangle \leq \int_0^T \big\langle \zeta, e^{A(T-s)} B w(s) \big\rangle ds = \sum_{i=1}^M \int_0^T \big\langle \zeta, e^{A(T-s)} b_i \big\rangle (u_i(s) - \bar{u}_i(s))\, ds, \tag{3.30}$$

where $B = \big(b_i\big)_{i=1,2,\ldots,M}$. Recalling (2.6) we have also

$$\langle \zeta, \bar{y} - \bar{x} \rangle = -\sum_{i=1}^M \int_0^T \big| \big\langle \zeta, e^{A(T-s) b_i} \big\rangle \big|\, \big| u_i(s) - \bar{u}_i(s) \big|\, ds. \tag{3.31}$$

Moreover,

$$\|\bar{y} - \bar{x}\| \leq e^{\|A\|T} \sum_{i=1}^{M} \|b_i\| \int_0^T |u_i(s) - \bar{u}_i(s)| ds. \tag{3.32}$$

We now apply the same argument leading to (3.28) to each summand of the right hand side of (3.30). Therefore we obtain, using (3.31), that

$$\langle \zeta, \bar{y} - \bar{x} \rangle \leq -C'(A, B, T, N, M) e^{-\|A\|T} \sum_{i=1}^{M} \Big(\int_0^T |u_i(s) - \bar{u}_i(s)| \, ds \Big)^N, \tag{3.33}$$

where the positive constant C' depends only on A, B, T, N, M and

$$\liminf_{T \to 0} C'(A, B, T, N, M) > 0.$$

We conclude the proof of (3.19) by applying (3.32) and setting

$$\gamma = 2^N e^{-(N+1)\|A\|T} C''(A, B, T, N, M),$$

where C'' is a constant enjoying the same properties as C'.

In order to prove the statement concerning the ball contained in $\mathcal{R}^T$, observe that the inequality (3.33) with $u \equiv 0$ and $\bar{y} = 0$, taking into account that the control $\bar{u}$ is necessarily bang-bang, becomes

$$\langle \zeta, \bar{x} \rangle \geq C'(A, B, T, N, M) e^{-\|A\|T} M T^N,$$

from which (taking $\|\zeta\| = 1$) we obtain

$$\|\bar{x}\| \geq C'(A, B, T, N, M) e^{-\|A\|T} M T^N := \gamma' T^N. \tag{3.34}$$

The above inequality yields in particular that 0 belongs to the interior of $\mathcal{R}^T$. Since $\mathcal{R}^T$ is convex and (3.34) holds for all $\bar{x} \in \operatorname{bdry} \mathcal{R}^T$, (3.20) follows.
The last statement follows from (3.29) and the explicit expressions for γ and γ'. The proof is concluded. □

Remark 3.4. *The exponent N in (3.19) is optimal.*

In fact, consider the dynamics

$$x^{(N)} = u, \qquad u \in [-1, 1].$$

Let $x_1(\cdot)$ be the solution corresponding to the control $u \equiv 1$. Fix $s > 0$ and let $x_s(\cdot)$ be the solution corresponding to the control

$$u_s(t) = \begin{cases} 1 & 0 < t < s \\ -1 & s < t. \end{cases}$$

Fix any $T > 0$ and observe that $x_1(T) = T^N/N!$, while, for $0 < s < T$,

$$x_s(T) - x_1(T) = \frac{-2}{N!}(T - s)^N.$$

Observe also that $|x_s^{(N-1)}(T) - x_1^{(N-1)}(T)| = 2|T - s|$, so that setting $X_i = (x_i, x_i', \ldots, x_i^{(N-1)})$, $i = 1, s$, one obtains

$$\langle e_1, X_s(T) - X_1(T) \rangle = \frac{-2}{N!}(T - s)^N \geq -\gamma \big\| X_s(T) - X_1(T) \big\|^N,$$

for a suitable positive constant γ. If $s \to T$, then $X_s(T) \to X_1(T)$, and this shows that N is the smallest exponent allowed in (3.19). □

3.2. **Nonautonomous systems.**

The following Lemma is a first step for studying the reachable sets in the case of nonlinear control systems by using the linearization approach which we design in this paper. We will prove that under the rank condition (normality) at 0 of the linear nonautonomous control system (3.35), the strict convexity of the reachable sets is preserved up to a sufficiently small time, provided $A(t)$ and $B(t)$ are not too far from $A(0)$, $B(0)$. Sufficient conditions for the validity of the assumptions if $N = 2$ will be given below. In Section 4 sufficient conditions in the case where A and B come from a linearization around a trajectory will also be given.

Let $A : \mathbb{R}^+ \to \mathbb{M}_{N\times N}$ and $B : \mathbb{R}^+ \to \mathbb{M}_{N\times M}$, $1 \le M \le N$, be measurable and consider the linear nonautonomous control system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \tag{3.35}$$

together with

$$\dot{x}(t) = A(0)x(t) + B(0)u(t), \tag{3.36}$$

where $u = (u_1, \dots, u_M) \in [-1, 1]^M$. We denote by $M^\top(\cdot, \cdot)$ the matrix solution of

$$\begin{cases} \frac{\partial}{\partial t}X(t,s) = X(t,s)A(t) \text{ for } t, s \ge 0 \\ M(s,s) = \mathbb{I} \end{cases} \tag{3.37}$$

and by $M_0^\top(\cdot, \cdot)$ the matrix solution of

$$\begin{cases} \frac{\partial}{\partial t}X(t,s) = X(t,s)A(0) \text{ for } t, s \ge 0 \\ M_0(s,s) = \mathbb{I}. \end{cases} \tag{3.38}$$

Let $b(\cdot)$ be a column of $B(\cdot)$, let $\mathcal{T} > 0$ and define, for any $\zeta \in \mathbb{R}^N$,

$$g(t) = \langle \zeta M^\top(\mathcal{T}, t), b(t)\rangle, \tag{3.39}$$

and

$$g_0(t) = \langle \zeta M_0^\top(\mathcal{T}, t), b(0)\rangle. \tag{3.40}$$

Observe that g and g_0 can be seen as the switching functions related to (3.35) and (3.36), respectively.

We state now an abstract result which permits to transfer to g some properties of g_0 and to establish the quantitative strict convexity estimate for the reachable set from the origin of (3.35). Sufficient conditions in order to apply the following lemma to suitable linearizations in dimension 2 and 3 will be given in Section 4.

Lemma 3.5. *Fix $\zeta \in \mathbb{R}^N$, $\|\zeta\| = 1$, and let g and g_0 be defined according to (3.39) and* (3.40)*, respectively. Assume that*

(i) $\operatorname{rank}[b(0), A(0)b(0), \dots, A^{N-1}(0)b(0)] = N$*;*

(ii) *g is $N-2$ times differentiable and $g^{(N-2)}$ is absolutely continuous;*

(iii) *there exists a constant $K = K(A, B)$ and a time $\mathcal{T}'$ such that for all $i = 0, \dots, N-1$ one has*

$$|g^{(i)}(t) - g_0^{(i)}(t)| \le Kt \qquad \text{for all } \ 0 \le t \le \mathcal{T}'.$$

Let $\mathcal{L}$ be defined by (3.4), with $A(0)$, $b(0)$ in place of A, b, respectively.
Then there exists $\mathcal{T} = \mathcal{T}(A, B, N, \mathcal{T}')$ with the following property:
for every $0 < \tau < \mathcal{T}$ there exist disjoint sets I_i and numbers $\mathcal{N}_i$, $i = 0, \dots, N-1$, depending only on $A(0), B(0), \tau$, such that

(a) $[0, \mathcal{T}] = \bigcup_{i=0}^{N-1} I_i$;

(b) each I_i, $i = 0, \dots, N-1$ is the disjoint union of at most $\mathcal{N}_i$ intervals;

(c) for each $i = 0, \dots, N-2$ and each $s \in I_i$

$$|g^{(i)}(s)| \geq \frac{\mathcal{L}}{2N} e^{-\|A(0)\|s};$$

(d) $g^{(N-1)}$ has constant sign in every connected component of I_{N-1} and, for each $s \in I_{N-1}$

$$|g^{(N-1)}(s)| \geq \frac{\mathcal{L}}{2N} e^{-\|A(0)\|s}.$$

Remark 3.6. *$I_0, \dots, I_{N-1}$ are exactly the intervals constructed in Lemma 3.2 with g_0 in place of g.*

Proof of Lemma 3.5. Let $I_0, \dots, I_{N-1}$ be the sets appearing in the statement of Lemma 3.2, with g_0 in place of g, so that, in particular, (a) and (b) hold. Moreover, for each $s \in I_i$ we have

$$|g_0^{(i)}(s)| \geq \frac{\mathcal{L}}{N} e^{-\|A(0)\|\tau}.$$

Therefore, if $K\mathcal{T} < \frac{\mathcal{L}}{2N} e^{-\|A(0)\|s}$, also (c) and (d) hold, owing to (iii). □

Remark 3.7. *Sufficient conditions for the validity of the assumptions of Lemma 3.5 in the case $N = 2$.*

Let $N = 2$. Sufficient conditions for the validity of assumptions (ii) and (iii) in the above Lemma are the following:

(C_0) $A(\cdot)$ is measurable and

$$\|A(t) - A(0)\| \leq Lt \text{ for all } t \geq 0;$$

(C_1) $B(\cdot)$ is absolutely continuous and

$$\left\| \frac{d}{dt} B(t) \right\| \leq 2Lt \text{ for all } t \geq 0,$$

where L is a positive constant.

Observe that condition (C_0) implies that $t = 0$ is a continuity point for $A(\cdot)$, so that $A(0)$ in (C_0) is meaningful. □

As an immediate corollary of Lemma 3.5 we obtain the following

Theorem 3.8. *Consider the linear nonautonomous control system (3.35) under the assumptions of Lemma (3.5). Let $\mathcal{R}^\tau$ denote the reachable set at time $\tau > 0$ from the origin for (3.35). Then there exist a time $\mathcal{T} = \mathcal{T}(A, B, N) > 0$ and a constant $\gamma = \gamma(A, B, N) > 0$ such that for every $0 \leq \tau \leq \mathcal{T}$, for every $x, y \in \mathcal{R}^\tau$, for every $\zeta \in N_{\mathcal{R}^\tau}(x)$, we have*

$$\langle \zeta, y - x \rangle \leq -\gamma \|\zeta\| \|y - x\|^N.$$

Moreover, there exists another constant $\gamma' = \gamma'(A, B, N) > 0$ such that for every $0 < \tau \leq \mathcal{T}$

$$\text{the ball } B(0, \gamma' \tau^N) \text{ is contained in } \mathcal{R}^\tau. \tag{3.41}$$

Proof. The argument developed in the proof of Theorem 3.3 can be used also in this case. Indeed, fix $\bar{x} \in \operatorname{bdry} \mathcal{R}^\tau$ together with a control $\bar{u}(\cdot)$ steering 0 to $\bar{x}$ in time τ and let $\zeta \in N_{\mathcal{R}^\tau}(\bar{x})$, $\|\zeta\| = 1$, be such that

$$\bar{u}(t) = \operatorname{sign}\langle \zeta, b(t)M(\tau,t)\rangle \qquad \text{for a.e. } t \in (0,\tau) \tag{3.42}$$

(here, as at the beginning of the proof of Theorem 3.3, we assume that $B = b$ is a vector, i.e., the control is scalar).
Then the proof proceeds exactly as for Theorem 3.3, provided that g is given by

$$g(t) = \langle \zeta, b(t)M(\tau,t)\rangle,$$

so that for all $\bar{y} \in \mathcal{R}^\tau$ one has

$$\langle \zeta, \bar{y} - \bar{x}\rangle = -2\sum_{i=0}^{N-1} \int_{I_i} |g(s)| K_1(s)\, ds,$$

where $K_1(s) = \frac{1}{2}|u(\tau - s) - \bar{u}(\tau - s)|$, $u(\cdot)$ is the control which steers the origin to $\bar{y}$, and the sets I_i are those appearing in the statement of Lemma 3.5. □

4. A nonlinear bang bang principle in dimensions 2 and 3

Starting from the present section we will deal with nonlinear control systems, which are affine and symmetric with respect to the control. This section is devoted to giving sufficient conditions so that controls steering the origin to the boundary of the reachable set are always bang-bang, provided the final time is sufficiently small. More precisely, the following result holds.

Theorem 4.1. *Let $N = 2$ or $N = 3$. Consider the control system*

$$\begin{cases} \dot{x}(t) = & F(x(t)) + G(x(t))u(t), \\ x(0) = & 0, \end{cases} \tag{4.1}$$

where $1 \le M \le N$, $u(\cdot) = (u_1(\cdot), \dots, u_M(\cdot)) \in [-1,1]^M$ a.e., and $F : \mathbb{R}^N \to \mathbb{R}^N$ and $G = (G_1, \dots, G_M) : \mathbb{R}^N \to \mathbb{M}_{N\times M}$ satisfy the following assumptions:

(i) *F and G are of class $\mathcal{C}^{N-1}$ and all partial derivatives are Lipschitz with constant L;*
(ii) *$F(0) = 0$;*
(iii) *$\operatorname{rank}[G_i(0), DF(0)G_i(0), \dots, (DF(0))^{N-1}G_i(0)] = N$ for $i = 1, \dots, M$;*
(iv) *if $N = 3$, then $DG(0) = 0$ and $D^2F(0) = 0$.*

Let $\mathcal{R}^\tau$ denote the reachable set of (4.1) at time $\tau > 0$. Then there exists $\mathcal{T} > 0$, depending only on $DF(0), G(0), L, N$, such that for every $0 < \tau < \mathcal{T}$ the following properties hold:

(a) *every admissible control $u(\cdot)$ such that the corresponding trajectory $y^u(\cdot)$ of (4.1) at time τ belongs to the boundary of $\mathcal{R}^\tau$ is essentially determined by the curve $\lambda(\cdot)$, the solution of the adjoint equation*

$$\begin{cases} \dot{\lambda}(t) & = -\lambda(t)\big(DF(y^u(t)) + DG(y^u(t))u(t)\big) \\ \lambda(\tau) & = \zeta \in N_{\mathcal{R}^\tau}(y^u(\tau)), \quad \|\zeta\| = 1, \end{cases} \tag{4.2}$$

through the identity

$$u(t) = \operatorname{sign}\langle \lambda(t), G(y^u(t))\rangle \quad \text{a.e.;} \tag{4.3}$$

(b) *u is bang-bang, i.e., $u(t) \in \{-1,1\}^M$ a.e., and is piecewise constant;*

(c) *the maximum number of discontinuities of u depends only on $DF(0)$, $G(0)$, L, and N.*

Proof. We consider first the case where $M = 1$, i.e., $G(x)$ is a vector and the control u is scalar. Fix $\tau > 0$ and an admissible control u such that $\bar{x} := y^u(\tau) \in \operatorname{bdry} \mathcal{R}^\tau$. By the Maximum Principle (see Theorem 2.6) there exist $\zeta \in N_{\mathcal{R}^\tau}(\bar{x})$, $\|\zeta\| = 1$, and an adjoint curve $\lambda(\cdot)$, a solution of (4.2), such that u satisfies (4.3). Proving (a), (b), and (c) amounts to showing that the switching function $\langle \lambda(t), G(y^u(t))\rangle$ vanishes at most finitely many times in $[0, \tau]$ and the number of its zeros depends only on $DF(0)$, $G(0)$, L, and N.
For convenience we rewrite the switching function in the following way. Let $M^\top(\cdot,\cdot)$ denote the matrix solution of (3.37), with $A(t) = DF(y^u(t)) + DG(y^u(t))u(t)$, and set $b(t) = G^\top(y^u(t))$, $t \in [0, \tau]$. Then

$$g(t) := \langle \lambda(t), b(t)\rangle = \langle \zeta, b(t)M(\tau, t)\rangle.$$

Let also $M_0^\top(\cdot,\cdot)$ be defined by (3.38), where $A(0) = DF(0)$, and let g_0 be defined according to (3.40). We wish to apply Lemma 3.5 to the above introduced mappings g and g_0.
First of all, we compute g' and observe that it is continuous. Indeed,

$$\begin{aligned}
g'(t) &= \big\langle \zeta, b'(t)M(\tau,t)\big\rangle + \big\langle \zeta, b(t)\frac{\partial}{\partial t}M(\tau,t)\big\rangle \\
&= \big\langle \zeta, DG^\top(y^u(t))\dot{y}^u(t)M(\tau,t)\big\rangle + \big\langle \zeta, -b(t)A^\top(t)M(\tau,t)\big\rangle \\
&= \big\langle \zeta, DG^\top(y^u(t))\big(F(y^u(t)) + G(y^u(t))u(t)\big)^\top M(\tau,t)\big\rangle \\
&\qquad + \big\langle \zeta, -G^\top(y^u(t))\big(DF(y^u(t)) + DG(y^u(t))u(t)\big)^\top M(\tau,t)\big\rangle \\
&= \big\langle \zeta, \big(DG^\top(y^u(t))F^\top(y^u(t)) - DF^\top(y^u(t))G^\top(y^u(t))\big)M(\tau,t)\big\rangle \\
&= \big\langle \lambda(t), [F^\top, G^\top](y^u(t))\big\rangle,
\end{aligned}$$

where $[F^\top, G^\top](x) := DG^\top(x)F^\top(x) - DF^\top(x)G^\top(x)$ denotes the Lie bracket.
Moreover, if $N = 3$ g' is a.e. differentiable and we have

$$g''(t) = \big\langle \dot{\lambda}(t), [F^\top, G^\top](y^u(t))\big\rangle + \big\langle \lambda(t), \frac{d}{dt}[F^\top, G^\top](y^u(t))\big\rangle.$$

Finally, observing that $g_0'(t) = \big\langle \zeta, -DF^\top(0)G^\top(0)M_0(\tau,t)\big\rangle = \big\langle \zeta M_0^\top(\tau,t), [F^\top, G^\top](0)\big\rangle$, we have

$$\begin{aligned}
g'(t) - g_0'(t) &= \big\langle \zeta\big(M^\top(\tau,t) - M_0^\top(\tau,t)\big), [F^\top, G^\top](y^u(t))\big\rangle \\
&\qquad + \big\langle \zeta M_0^\top(\tau,t), DG^\top(y^u(t))F^\top(y^u(t))\big\rangle \\
&\qquad + \big\langle \zeta, \big(DF^\top(0) - DF^\top(y^u(t))\big)G^\top(y^u(t))\big\rangle \\
&\qquad + \big\langle \zeta, DF^\top(0)\big(G^\top(0) - G^\top(y^u(t))\big)\big\rangle \\
&= I + II + III + IV.
\end{aligned}$$

We are now going to estimate separately each summand of the above expression.
First, we observe that, thanks to the assumptions (i) and (ii), we have

$$\begin{aligned}
\|M(\tau,t) - M_0(\tau,t)\| &\le \int_0^t \Big(\|A(s)\|\,\|M_0(\tau,s) - M(\tau,s)\| + \|M_0(\tau,s)\|\,\|A(0) - A(s)\|\Big)\,ds \\
&\le K_1 \int_0^t \|M(\tau,s) - M_0(\tau,s)\|\,ds + K_2 t^2,
\end{aligned}$$

where K_1 and K_2 are suitable constants depending only on $DF(0)$, $G(0)$, L, and τ. Gronwall's lemma therefore yields

$$\|M(\tau,t) - M_0(\tau,t)\| \le K_3 t^2 \qquad \text{for all } \ t \in [0,\tau],$$

where the constant K_3 depends only on K_1, K_2. Therefore, there exists a constant K_I, depending only on $DF(0)$, $G(0)$, L, and τ such that

$$|I| \le K_I t^2 \qquad \text{for all } \ t \in [0,\tau]. \tag{4.4}$$

Assumptions (i) and (ii) yield in turn

$$|II| \le K_{II} t \qquad \text{for all } \ t \in [0,\tau], \tag{4.5}$$

for a suitable constant K_{II} depending only on $DF(0)$, L, and τ, and

$$|III| \le K_{III} t \qquad \text{for all } \ t \in [0,\tau], \tag{4.6}$$

$$|IV| \le K_{IV} t \qquad \text{for all } \ t \in [0,\tau], \tag{4.7}$$

where again K_{III} and K_{IV} depend only on $DF(0)$, $G(0)$, L, and τ.
Therefore, summing (4.4), (4.5), (4.6), and (4.7), we obtain that there exist K and $\mathcal{T}' > 0$, depending only on $DF(0)$, $G(0)$, and L, such that

$$|g'(t) - g_0'(t)| \le Kt \qquad \text{for all } \ 0 \le t \le \mathcal{T}'. \tag{4.8}$$

Let now $N = 3$, and observe that owing to assumption (iv) each summand I, II, III, IV, divided by t^2, is bounded and a.e. differentiable, so that

$$|g''(t) - g_0''(t)| \le K't \qquad \text{for a.e. } \ t \in [0,\mathcal{T}'],$$

where the constant K' depends only on $DF(0)$, $G(0)$, and L and so does $\mathcal{T}'$.
Observe that all the above constants do not depend on ζ. Therefore, invoking assumption (iii), we can apply Lemma 3.5 (for $N = 2, 3$), thus obtaining that there exists $\mathcal{T} > 0$, depending only on $G(0)$, $DF(0)$, L, and N with all the properties (a), (b), (c), and (d), which are exactly those required to complete the proof.
In the general case (i.e., $1 < M \le N$), it suffices to apply the above argument to each column of G. The proof is concluded. □

Remark 4.2. *On the assumption $N = 2$ or $N = 3$.*

The restriction $N \le 3$ depends on our method for comparing the switching function g for the nonautonomous system coming from the linearization along an optimal trajectory, with the switching function for autonomous system obtained by linearizing at the origin. This comparison requires higher order derivatives of g, whose existence we are not able to insure if $N > 3$. □

5. Quantitative strict convexity of reachable sets and uniqueness of optimal controls: the nonlinear two dimensional case

In this subsection, we will show that, provided the linearization at 0 satisfies the normality condition and the nonlinear part is smooth and small enough, the reachable set is strictly convex up to a sufficiently small time.

Theorem 5.1. *Consider the control system*

$$\begin{cases} \dot{x}(t) = F(x(t)) + G(x(t))u(t), \\ x(0) = 0, \end{cases} \tag{5.1}$$

under the following assumptions (in the following M is either 1 or 2):
$u(\cdot) = (u_1(\cdot), u_M(\cdot)) \in [-1,1]^M$ *a.e., $F : \mathbb{R}^2 \to \mathbb{R}^2$, $G : \mathbb{R}^2 \to \mathbb{M}_{2\times M}$ are of class $\mathcal{C}^{1,1}$ (with Lipschitz constant L) and*

(i) $F(0) = 0$,
(ii) Rank $[G_i(0), DF(0)G_i(0)] = 2$ *for* $i = 1, M$ *where* $G = (G_1, G_M)$,
(iii) $DG(0) = 0$.

Let $\mathcal{R}^\tau$ denote the reachable set at time $\tau > 0$ for (5.1). Then there exists $\mathcal{T} > 0$, depending only on $L, DF(0), G(0)$, with the following properties:

(a) *for every $\tau \le \mathcal{T}$ and every $x \in \operatorname{bdry} \mathcal{R}^\tau$ there exists one and only one admissible control u steering the origin to x in time τ (and u is bang-bang with finitely many switchings).*
(b) *For every $0 < \tau < \mathcal{T}$ the reachable set $\mathcal{R}^\tau$ is strictly convex. More precisely, for every $x_1 \in \operatorname{bdry} \mathcal{R}^\tau$ and $x_2 \in \mathcal{R}^\tau$, for every $\zeta \in N^P_{\mathcal{R}^\tau}(x_1)$, one has*

$$\langle \zeta, x_2 - x_1 \rangle \le -\gamma \|\zeta\| \|x_2 - x_1\|^2. \tag{5.2}$$

where γ is a positive constant depending only on L, $DF(0)$, $G(0)$.
(c) *There exist another time $\mathcal{T}' > \mathcal{T}$, depending only on L, $DF(0)$, $G(0)$, such that for every $0 < \tau < \mathcal{T}'$ the reachable set $\mathcal{R}^\tau$ has positive reach. More precisely, for every $x_1 \in \operatorname{bdry} \mathcal{R}^\tau$ and $x_2 \in \mathcal{R}^\tau$, for every $\zeta \in N^P_{\mathcal{R}^\tau}(x_1)$, one has*

$$\langle \zeta, x_2 - x_1 \rangle \le \gamma' \|\zeta\| \|x_2 - x_1\|^2, \tag{5.3}$$

where γ' is a nonnegative constant depending only on L, $DF(0)$, $G(0)$.
(d) *There exist another positive constant γ'' and a positive time $\mathcal{T}'' \le \mathcal{T}$, depending only on L, $DF(0)$, $G(0)$, such that the ball $B(0, \gamma'' \tau^2)$ is contained in $\mathcal{R}^\tau$ for all $0 < \tau < \mathcal{T}''$.*
(e) *The minimum time function is continuous in $\mathcal{R}^\tau$, for all $0 < \tau < \mathcal{T}''$.*

Proof. We begin proving the result for $M = 1$, i.e., for a scalar control.

Fix $\tau > 0$ and $x_1 \in \operatorname{bdry} \mathcal{R}^\tau$, together with an optimal control $u_1(\cdot)$ steering 0 to x_1 and the associate trajectory $x_1(\cdot)$. Take any $x_2 \in \mathcal{R}^\tau$ together with $u_2(\cdot)$ steering 0 to x_2 and the associate trajectory $x_2(\cdot)$, and set $x(t) = x_2(t) - x_1(t)$. Then, for a.e. $t \in [0, \tau]$,

$$\dot{x}(t) = A_1(t)x(t) + G(x_1(t))w(t), \tag{5.4}$$

where $w(t) = u_2(t) - u_1(t)$ and

$$A_1(t) = \int_0^1 DF(x_1(t) + \tau x(t))d\tau + u_2(t) \int_0^1 DG(x_1(t) + \tau x(t))d\tau.$$

Let $z(\cdot)$ be the solution of the linear system which is defined by linearizing along the optimal trajectory $x_1(\cdot)$:

$$\begin{cases} \dot{z}(t) = A(t)z(t) + G(x_1(t))w(t), \\ z(0) = 0, \end{cases} \tag{5.5}$$

where $A(t) = DF(x_1(t)) + DG(x_1(t))u_1(t)$.
We have

$$\begin{aligned} \frac{d}{dt}\|x(t) - z(t)\| &\le \|A_1(t)x(t) - A(t)z(t)\| \\ &\le \|A(t)\|\|x(t) - z(t)\| + \|A_1(t) - A(t)\|\|x(t)\| \\ &\le L_1\|x(t) - z(t)\| + L\|x(t)\|^2, \end{aligned}$$

where $L_1 = \|DF(0)\| + 2Le^{2L\tau}$.
Thus, by Gronwall's inequality we get

$$\|x(t) - z(t)\| \le e^{L_1 t} L \int_0^t \|x(s)\|^2 \, ds. \tag{5.6}$$

On the other hand, observing that

$$\frac{d}{dt}\|x(t)\| \le L_2\|x(t)\| + L_3|w(t)|,$$

(where $L_2 = \|DF(0)\| + 4Le^{2L\tau}$ and $L_3 = |G(0)| + e^{2L\tau}$) we also have

$$\|x(t)\| \le L_3 e^{L_2 t} \int_0^t |w(s)| ds. \tag{5.7}$$

From (5.6) and (5.7), one obtains

$$\|x(t) - z(t)\| \le L_4 t \Big(\int_0^t |w(s)| ds \Big)^2, \tag{5.8}$$

where $L_4 = LL_3^2 e^{(L_1+2L_2)\tau}$.

Since $\bar{x}_1 \in \operatorname{bdry} \mathcal{R}^\tau$, by Pontryagin's maximum principle there exists an absolutely continuous function $\lambda : [0, \tau] \to \mathbb{R}^2$ with the following properties

$$\dot{\lambda}(t) = -\lambda(t)A(t), \quad \lambda(\tau) = \zeta \in N_{\mathcal{R}^\tau}(\bar{x}_1),\ \zeta \neq 0,$$

$$u_1(t) = \operatorname{sign}\langle \lambda(t), G(x_1(t)) \rangle. \tag{5.9}$$

We set now $b(t) = G(x_1(t))$ and consider the linear nonautonomous control system

$$\begin{cases} \dot{y}(t) = A(t)y(t) + b(t)u(t), \\ y(0) = 0, \end{cases} \tag{5.10}$$

together with the trajectory $y_1(\cdot)$, corresponding to the control $u_1(\cdot)$. Observe first that, thanks to (5.9), x_1 belongs to the boundary of the reachable set at time τ for (5.10). Observe moreover that $A(\cdot)$ is measurable, and, since both F and G are Lipschitz with constant L and $DG(0) = 0$, we have

$$\begin{aligned} \|A(t) - A(0)\| &= \|DF(x_1(t)) + DG(x_1(t))u_1(t) - DF(0)\| \\ &\le 2L\|x_1(t)\| \le 2L\|G(0)\|e^{2Lt}t. \end{aligned}$$

Finally, $b'(t) = DG(x_1(t))\dot{x}_1(t)$ so that

$$\|b'(t) - b(0)\| = \|b'(t)\| \le L\|x_1(t)\|\big(2L\|x_1(t)\| + \|G(0)\|\big) \le Kt,$$

where $K = Le^{2L\tau}\big(2Le^{2L\tau} + 1\big)\|G(0)\|^2$.
Therefore, all assumptions of Theorem 3.8 are satisfied by (5.10), so that there exists a time $\mathcal{T}_0 > 0$ depending only on $L, DF(0), G(0)$ such that if $0 \le \tau < \mathcal{T}_0$ the following properties hold:

i) $u_1(t)$ is uniquely determined by (5.9) on $(0, \tau)$,
ii) there exists a constant $\gamma(\tau) > 0$, depending only on $L, DF(0), G(0), \tau$ such that, for all trajectories $y(\cdot)$ of (5.10), $\langle \zeta, y(\tau) - y_1(\tau) \rangle \le -\gamma(\tau)\|y(\tau) - y_1(\tau)\|^2$. More precisely, recalling (3.28),

$$\langle \zeta, y_2(\tau) - y_1(\tau) \rangle \le -\gamma_1(\tau) \Big(\int_0^\tau |w(s)| ds \Big)^2, \tag{5.11}$$

where $y_2(\cdot)$ is the trajectory of (5.10) associated with the control $u_2(\cdot)$ and γ_1 enjoys the same properties of γ.

We remark (see (3.29)) that $\gamma_1(\tau)$ is bounded away from 0 as $\tau \to 0^+$. Moreover, one can see that $z(t) = y_2(t) - y_1(t)$. Therefore

$$\begin{aligned} \langle \zeta, x_2 - x_1 \rangle &= \langle \zeta, x(\tau) - z(\tau) \rangle + \langle \zeta, z(\tau) \rangle \\ &\leq \|x(\tau) - z(\tau)\| + \langle \zeta, z(\tau) \rangle. \end{aligned} \tag{5.12}$$

Recalling (5.8) and (5.11), we obtain

$$\langle \zeta, x_2 - x_1 \rangle \leq (L_4 \tau - \gamma_1(\tau)) \Big(\int_0^\tau |w(s)| ds \Big)^2. \tag{5.13}$$

Thus if $\tau \leq \frac{1}{2L_4} \liminf_{\tau \to 0^+} \gamma(\tau) =: \mathcal{T}_1$ then

$$\langle \zeta, x_2 - x_1 \rangle \leq -\frac{\gamma_1(\tau)}{2} \Big(\int_0^\tau |w(s)| ds \Big)^2.$$

From this inequality the uniqueness of the control steering the origin to x_1 in time τ follows immediately by contradiction. Setting $\mathcal{T} = \min\{\mathcal{T}_0, \mathcal{T}_1\}$ and recalling (5.7) we obtain (5.2). The proof of the strict convexity is completed by applying Proposition 7.3.

The proof of (5.3) is entirely analogous, where it suffices to take $\mathcal{T}' = \mathcal{T}_0$.

We consider now the statement concerning the ball contained in the reachable set. To this aim, take $u_2 \equiv 0$ and set y_1 to be the solution of (5.5) with u_1 in place of w. Then (5.8) yields, for all $t > 0$,

$$\|x_1(t) - y_1(t)\| \leq L_4 t^3.$$

Recalling (3.41), we obtain from the previous inequality that

$$\|x_1(t)\| \geq \|y_1(t)\| - L_4 t^3 \geq \tilde{\gamma} t^2 - L_4 t^3,$$

for a suitable constant $\tilde{\gamma}$, which yields in particular that 0 belongs to the interior of $\mathcal{R}^t$, $0 \leq t \leq \mathcal{T}$. Since the above argument can be repeated for every point in the boundary of $\mathcal{R}^t$ and the constant L_4 is independent of the reference point, the statement follows by recalling that we already proved that $\mathcal{R}^t$ is convex for all $0 \leq t \leq \mathcal{T}$.

The continuity of T follows easily from the the fact that reachable sets contain a ball (see, e.g., Propositions IV.1.2 and IV.1.6 in [1]).

In the case $M = 2$, it suffices to apply the above arguments to each control. □

The following Remark follows immediately from the proof of Theorem 5.1.

Remark 5.2. *Let $x_1(\cdot)$ and $\lambda(\cdot)$ be as in the proof of Theorem 5.1. For all $0 < t \leq \tau = T(x_1)$, one has $\lambda(t) \in N_{\mathcal{R}^t}(x_1(T(x_1 - t)))$. More precisely,*

$$\langle \lambda(t), y - x_1(T(x_1) - t) \rangle \leq -\gamma \|\lambda(t)\| \|y - x_1(T(x_1) - t)\|^2$$

for all $y \in \mathcal{R}^t$, where γ is the constant appearing in (5.2).

In fact, put $\lambda(t)$ in place of ζ in (5.12). Then the first summand can be estimated in the same way, while the upper bound on the second summand, namely the analogue of (5.11), can be obtained through the same arguments leading to (5.11). □

We conclude this section with a counterexample showing the sharpness of assumption (iii) in Theorem 5.1.

Remark 5.3. *An example of a two dimensional nonlinear control system satisfying assumptions (i) and (ii) of Theorem 5.1 such that the reachable set $\mathcal{R}^\tau$ is not convex for all $\tau > 0$.*

Consider the control system

$$\begin{cases} \dot{x}_1 &= x_2(1+u) \\ \dot{x}_2 &= u, \qquad u \in [-1,1]. \end{cases} \tag{5.14}$$

Set $F(x_1,x_2) = (x_2,0)$, $G(x_1,x_2) = (x_2,1)$, and observe that the assumptions (i) and (ii) of Theorem 5.1 are satisfied, while (iii) is not. The Hamiltonian for this system is

$$\mathcal{H}\big((x_1,x_2),(p_1,p_2),u\big) = x_2p_1 + u(p_1x_2+p_2),$$

and Pontryagin's Maximum Principle states that if $\bar{x}(\cdot)$ is an optimal trajectory corresponding to the control $\bar{u}(\cdot)$, then there exists a function $\lambda = (\lambda_1,\lambda_2)$, never vanishing, and a constant $\lambda_0 \le 0$ such that, for a.e. t,

$$\begin{aligned} \lambda_1(t) &\equiv \lambda_1 \neq 0 \\ \dot{\lambda}_2(t) &= -(1+\bar{u}(t))\lambda_1 \\ \bar{x}_2(t)\lambda_1 + \bar{u}(t)(\lambda_1\bar{x}_2(t)+\lambda_2(t)) + \lambda_0 &= 0 \\ \bar{u}(t)(\lambda_1\bar{x}_2(t)+\lambda_2(t)) &= \max_{|u|\le 1} u(\lambda_1\bar{x}_2(t)+\lambda_2(t)). \end{aligned}$$

Now, observe that

$$\begin{aligned} \lambda_1\bar{x}_2(t)+\lambda_2(t) &= \lambda_1\int_0^t \bar{u}(s)\,ds - \lambda_1 t - \lambda_1\int_0^t \bar{u}(s)\,ds + \lambda_2(0) \\ &= -\lambda_1 t + \lambda_2(0). \end{aligned}$$

Therefore the function $\lambda_1\bar{x}_2(t)+\lambda_2(t)$ has at most one zero, so that the optimal control $\bar{u}$ is unique, bang-bang, and has at most one switching.
Fix now $\tau > 0$ and $0 < s < 1$ and consider the control

$$u_s(t) = \begin{cases} 1 & 0 < t < s\tau \\ -1 & s\tau < t < \tau. \end{cases}$$

The trajectory of (5.14) emanating from the origin and corresponding to the control u_2 is, at time τ,

$$\bar{x}^1_s := x^1_s(\tau) = s^2\tau^2, \quad \bar{x}^2_s := x^2_s(\tau) = \tau(2s-1).$$

Simple computations show that any other bang-bang control with at most one switching cannot reach $(\bar{x}^1_s,\bar{x}^2_s)$ at a time $\tau' < \tau$, for all $0 < s < 1$. Thus u_s is optimal. In particular, the curve $\gamma(s) := (\bar{x}^1_s,\bar{x}^2_s)$ belongs to the boundary of the reachable set $\mathcal{R}^\tau$. The unique unit normal to the curve $\gamma(s)$ at $s = 1/2$ which points outside $\mathcal{R}^\tau$ is

$$\zeta = \frac{(2,-\tau)}{\sqrt{4+\tau^2}}.$$

We compute:

$$\big\langle \zeta, \gamma(s) - \gamma(\frac{1}{2})\big\rangle = \frac{2\tau^2(s-\frac{1}{2})^2}{\sqrt{4+\tau^2}},$$

which implies that $\mathcal{R}^\tau$ is not convex. Observe however that $\mathcal{R}^\tau$ has positive reach at $\gamma(\frac{1}{2})$ for every $\tau > 0$, since

$$\langle \zeta, \gamma(s) - \gamma(\frac{1}{2}) \rangle \leq \frac{8}{(16+\tau^4)\sqrt{4+\tau^2}} \left\| \gamma(s) - \gamma(\frac{1}{2}) \right\|^2.$$

□

Remark 5.4. *On the assumption $N = 2$.*

Motivations for the restriction $N = 2$ are twofold. First of all, our analysis is based on the switching function of the nonautonomous system obtained by linearizing around an optimal trajectory, and this method requires $N \leq 3$ (see Remark 4.2). Second, the distance between trajectories of the nonlinear system (5.1) and of the linearized system (5.5) is of order two with respect to the control (see (5.8)), and this quadratic perturbation can be balanced by the strict convexity of the reachable set of the linearized system only if $N = 2$ (see (5.13)). □

6. Further results for the nonlinear two dimensional case

This section is devoted to proving that the epigraph of the minimum time function has positive reach, under the assumptions of Theorem 5.1. To this aim, a results of optimal points, i.e., on points which are *crossed* by an optimal trajectory, is needed.

6.1. **Optimal points.**

The classical definition of optimal point reads as follows.

Definition 6.1. *Let $x \in \mathbb{R}^N \setminus \{0\}$. We say that x is optimal if and only if there exists a point x_1 such that $T(x_1) > T(x)$ and a control u with the property that $y^{x_1,u}(\cdot)$ steers x_1 to 0 in the optimal time $T(x_1)$ and $x = y^{x_1,u}(T(x_1) - T(x))$.*

The following is the result on optimal points which will be used in the next subsection in order to ensure the positive reach of the epigraph of the minimum time function. It is based on the same estimates which lead to the strict convexity of the reachable set, and so it is restricted to two dimensional control systems.

Theorem 6.2. *Let $N = 2$ and let the assumptions of Theorem 5.1 be satisfied. Let $\mathcal{T} > 0$ be such that, according to Theorem 5.1, for all $0 \leq \tau < \mathcal{T}$, the reachable set $\mathcal{R}^\tau$ of (5.1) satisfies (5.2) for all $0 < \tau < \mathcal{T}$. Let $\bar{x}$ be such that $T(\bar{x}) < \mathcal{T}$. Then $\bar{x}$ is an optimal point.*

Proof. We consider first the case where G is a vector and the control u is one-dimensional. Set $\tau = T(\bar{x})$ and let $\bar{u}(\cdot)$ be the admissible control steering $\bar{x}$ to 0 in the optimal time τ, together with the associate trajectory $\bar{x}(\cdot)$. Set, for all $t \in [0, \tau]$,

$$A(t) = DF(\bar{x}(t)) + DG(\bar{x}(t))\bar{u}(t), \qquad b(t) = G(\bar{x}(t))$$

We assume preliminarily that $\bar{x}$ belongs to the boundary of $\mathcal{R}^\tau$ and let, by the Maximum Principle, λ be a solution of

$$\begin{cases} \dot{\lambda}(t) = -\lambda(t)A(t), \\ \lambda(\tau) = \zeta, \end{cases} \tag{6.1}$$

where $\zeta \in N_{\mathcal{R}^\tau}(\bar{x})$, $\zeta \neq 0$, and, for a.e. $t \in [0, \tau]$,

$$\bar{u}(t) = \operatorname{sign}\langle \lambda(t), b(t) \rangle. \tag{6.2}$$

Set, for $t \in [0, \tau]$,

$$g(t) = \langle \lambda(t), b(t) \rangle.$$

We are now going to extend $\bar{u}(\cdot)$ in an interval $[\tau, \tau + \delta]$ for a suitable $\delta > 0$, with the property that the extended control and its associate trajectory satisfy the Maximum Principle.

Three cases may occur:

(i) $g(\tau) > 0$,
(ii) $g(\tau) < 0$,
(iii) $g(\tau) = 0$.

In the first case, we set $\bar{u}(t) = 1$ for all $t > \tau$ and let $\bar{x}(\cdot)$ be the associate trajectory satisfying $\bar{x}(\tau) = \bar{x}$. We extend analogously $A(\cdot)$, $b(\cdot)$, $\lambda(\cdot)$ and $g(\cdot)$ for $t > \tau$. Set $g(\tau) := \delta_1$. Observe that g is locally Lipschitz, so that, for $t > \tau$,

$$g(t) = g(\tau) + g(t) - g(\tau) > \delta_1 - L_1(t - \tau)$$

for a suitable constant L_1. Therefore we can find $\delta > 0$ such that $0 \leq \tau + \delta < \mathcal{T}$ and $g(t) > 0$ for all $t \in [\tau, \tau + \delta]$, i.e.,

$$\bar{u}(t) = \text{sign } g(t) \quad \forall t \in [\tau, \tau + \delta].$$

The second case is entirely analogous, by substituting 1 with -1.
We consider now the third case. Let the I_0, I_1 be given by Lemma 3.5 for the function g in the interval $[0, \tau]$. Observe that necessarily $\tau \in I_1$, so that, in particular, $g'(\tau) \neq 0$. We set, for $t > \tau$

$$\bar{u}(t) = 1 \quad \text{if} \quad g'(\tau) > 0$$

or

$$\bar{u}(t) = -1 \quad \text{if} \quad g'(\tau) < 0$$

and let $\bar{x}(\cdot)$ be the associate trajectory satisfying $\bar{x}(\tau) = \bar{x}$. We extend analogously $A(\cdot), b(\cdot), \lambda(\cdot)$ and $g(\cdot)$ for $t > \tau$. Observe that

$$\dot{g}(t) = \langle \lambda(t), [F, G](\bar{x}(t)) \rangle,$$

where $[F, G](x) = DG(x)F(x) - DF(x)G(x)$ denotes the Lie bracket of F and G. Therefore $\dot{g}$ is continuous, so that there exists $\delta > 0$ such that the sign of $\dot{g}(t)$ equals the sign of $\dot{g}(\tau)$ for all $t \in [\tau - \delta, \tau + \delta]$. Therefore our construction of $\bar{u}(\cdot)$ on $[0, \tau + \delta]$ is such that for a.e. $t \in [0, \tau + \delta]$,

$$\bar{u}(t) = \text{sign } g(t).$$

Consequently, all conclusions of Theorem 5.1 hold up to the time $\tau + \delta$. In particular, for all $t \in [0, \tau + \delta]$, $\bar{x}(t) \in \text{bdry}\, \mathcal{R}^t$. Since $T(\cdot)$ is continuous in a neighborhood of the trajectory $\bar{x}(\cdot)$, we obtain that $\bar{u}(\cdot)$ steers the origin optimally to $\bar{x}(\tau + \delta)$ in time $\tau + \delta$. Since the above argument can be applied also to the reversed dynamics $\dot{x} = -F(x) + G(x)u$, $u \in [-1, 1]$, then $T(\bar{x}(\tau + \delta)) = \tau + \delta$. Let us now drop the assumption $\bar{x} \in \text{bdry}\, \mathcal{R}^\tau$. Since T is strictly decreasing along the optimal trajectory $\bar{x}(\cdot)$, and so $\bar{x}(t) \in \text{bdry}\, \mathcal{R}^{\tau - t}$ for all $0 < t < \tau$, there exists a nontrivial adjoint vector $\lambda(\cdot)$ which uniquely determines $\bar{u}(t)$ as in (6.2) up to the time τ. Thus the above argument can be applied also to $\bar{x}$.
If G is a 2×2 matrix, it suffices to perform the above construction for each column of G. The proof is concluded. □

We conclude the section with two corollaries. The first one is an immediate consequence of the proof of Theorem 6.2.

Corollary 6.3. *Under the same assumptions of Theorem 6.2, let $\tau = T(\bar{x}) < \tau_1 < \mathcal{T}$. Then there exists $x_1 \in \operatorname{bdry} \mathcal{R}^{\tau_1}$ and a control $u_1 : [\tau, \tau_1] \to [-1, 1]$ such that the trajectory $\tilde{x}(\cdot)$ corresponding to the control*

$$\tilde{u}(t) = \begin{cases} \bar{u}(t) & 0 \le t \le \tau, \\ u_1(t) & \tau < t \le \tau_1. \end{cases}$$

and such that $\tilde{x}(0) = x_1$ reaches 0 in the optimal time τ_1 and moreover $\tilde{x}(\tau_1 - \tau) = \bar{x}$.

From Corollary 6.3 we obtain that $\bar{x}(\cdot)$ and $\lambda(\cdot)$ in the proof of Theorem 6.2 can be extended up to the time $\mathcal{T}$. Therefore, we obtain the following further corollary.

Corollary 6.4. *Under the assumptions of Theorem 5.1, the maximized Hamiltonian along $\bar{x}(\cdot)$ associated with $\lambda(\cdot)$ is constant in $[0, \mathcal{T})$, i.e.,*

$$H(\bar{x}(t), \lambda(t)) = C \quad \forall t \in [0, \mathcal{T}).$$

Proof. Let G be a vector and so the control u be scalar. Then

$$H(\bar{x}(t), \lambda(t)) = \langle \lambda(t), F(\bar{x}(t)) \rangle + |\langle \lambda(t), G(\bar{x}(t)) \rangle|.$$

Observe that the switching function $g(t) = \langle \lambda(t), G(\bar{x}(t)) \rangle$ vanishes at most on a countable subset of $[0, \mathcal{T})$. Therefore, for a.e. $t \in [0, \mathcal{T})$, we have

$$\frac{d}{dt} H(\bar{x}(t), \lambda(t)) = \Big(- \langle \lambda(t), [F, G](\bar{x}(t)) \rangle + \langle \lambda(t), [F, G](\bar{x}(t)) \rangle \Big) \operatorname{sign} g(t) = 0.$$

If G is a 2×2 matrix, it suffices to perform the above computation for each column. The proof is concluded. □

6.2. **The epigraph of the minimum time function has positive reach.**
The present section is devoted to studying the "convexity type" of the minimum time function $T(\cdot)$, in the case where the dynamics satisfies a weak controllability condition, i.e., the function $T(\cdot)$ is merely continuous. The statement is two dimensional, since it is based on Theorem 6.2.

Theorem 6.5. *Let $N = 2$ and let the assumptions of Theorem 5.1 hold. Let $\mathcal{T}$ be given by Theorem 5.1. Then for every $0 < \tau < \mathcal{T}$ the epigraph of the minimum time function $T(\cdot)$ on $\mathcal{R}^\tau$ has positive reach.*

Corollary 6.6. *Under the same assumptions of Theorem 6.5 the minimum time function T satisfies all the properties listed in Theorem 2.4.*

Before beginning the proof of Theorem 6.5 we introduce the minimized Hamiltonian and study its sign.

Definition 6.7. *Let $x, \zeta \in \mathbb{R}^N$. We define the minimized Hamiltonian for the control system in* (5.1) *as*

$$h(x, \zeta) = \langle \zeta, F(x) \rangle + \min_{u \in \mathcal{U}} \langle \zeta, G(x) u \rangle.$$

Proposition 6.8. *Let x belong to the boundary of the reachable set $\mathcal{R}^\tau$ for* (5.1) *for some $\tau > 0$. Let $\zeta \in N^F_{\mathcal{R}^\tau}(x)$[1]. Then $h(x, \zeta) \le 0$.*

[1]here $N^F_{\mathcal{R}^\tau}(x)$ denotes the Fréchet normal cone to $\mathcal{R}^\tau$ at x, i.e., all vectors v such that $\limsup_{\mathcal{R}^\tau \ni y \to x} \langle v, (y - x)/\|y - x\| \rangle \le 0$

Proof. Let $\bar u(\cdot)$ be an admissible control steering x to 0 in time τ, together with the associate trajectory $\bar x(\cdot)$. Then, for all $0 \le t \le \tau$ the point $\bar x(t)$ belongs to $\mathcal{R}^\tau$, so that, by definition of Fréchet normal we have

$$\limsup_{t\to 0^+} \Big\langle \zeta, \frac{\bar x(t)-x}{\|\bar x(t)-x\|}\Big\rangle \le 0.$$

Observing that $\|x(t)-x\| \le Kt$ for a suitable constant K, we have

$$\limsup_{t\to 0} \Big\langle \zeta, \frac{\bar x(t)-x}{t}\Big\rangle \le 0.$$

In other words,

$$\begin{aligned} 0 &\ge \limsup_{t\to 0}\Big\langle \zeta, \frac{1}{t}\int_0^t \big(F(\bar x(s))+G(\bar x(s))\bar u(s)\big)ds\Big\rangle \\ &= \langle \zeta, F(x)\rangle + \limsup_{t\to 0}\Big\langle \zeta, G(x)\frac{\int_0^t \bar u(s)ds}{t}\Big\rangle. \end{aligned}$$

Let $t_n \to 0$ be a sequence such that $\lim_{n\to\infty}\frac{1}{t_n}\int_0^{t_n}\bar u(s)ds := \tilde u$ exists. By the convexity of $\mathcal{U}$, $\tilde u \in \mathcal{U}$, and so $h(x,\zeta) \le \langle \zeta, F(x)\rangle + \langle \zeta, G(x)\tilde u\rangle \le 0$. □

We are now ready to prove Theorem 6.5.

Proof of Theorem 6.5. Let $x \ne 0$ be such that $T(x) < \mathcal{T}$ and let $(\bar u(\cdot), \bar x(\cdot))$ be an optimal pair for x. By Maximum Principle[2] there exists $0 \ne \zeta \in N_{\mathcal{R}^{T(x)}}(x)$ such that the adjoint arc λ, with

$$\begin{cases} \dot\lambda(t) = \lambda(t)\big(DF(\bar x(t)) + DG(\bar x(t))\bar u(t)\big), \\ \lambda(T(x)) = \zeta \end{cases}$$

satisfies

$$\big\langle \lambda(t), F(\bar x(t)) + G(\bar x(t))\bar u(t)\big\rangle = h(\bar x(t),\lambda) \quad \text{for a.e. } t \in (0, T(x)).$$

We claim that

$$(\zeta, h(x,\zeta)) \in N^P_{\mathrm{epi}(T)}(x, T(x)), \tag{6.3}$$

i.e., there exists a constant $\sigma > 0$ such that, for all $y \in \mathbb{R}^N$ with $0 < T(y) < \mathcal{T}$ and for all $\beta \ge T(y)$, we have

$$\big\langle (\zeta,\theta), (y,\beta) - (x, T(x))\big\rangle \le \sigma\|(\zeta,\theta)\|\Big(\|y-x\|^2 + |\beta - T(x)|\Big), \tag{6.4}$$

where $\theta = h(x,\zeta)$, and, moreover,

$$\sigma \text{ is independent of } x \text{ and } \zeta. \tag{6.5}$$

Indeed, we consider two cases:

(a) $T(y) \le T(x)$;
(b) $T(y) > T(x)$.

In the first case, $y \in \mathcal{R}^{T(x)}$, so that by Theorem 5.1

$$\langle \zeta, y - x\rangle \le 0.$$

If $\beta \ge T(x)$ then (6.4) is automatically satisfied, since $\theta \le 0$ by Proposition 6.8. If instead $\beta < T(x)$, we set $x_1 = \bar x(\beta)$.

[2]observe that we are applying Theorem 2.6 for the *reversed dynamics* $\dot x = -F(x) - G(x)u$

We estimate first $\langle \zeta, y - x_1 \rangle$. Since $y \in \mathcal{R}^\beta$, recalling Remark 5.2, we have for suitable constants K_1, K_2 given by Gronwall's Lemma,

$$\begin{aligned}
\langle \zeta, y - x_1 \rangle &= \langle \lambda(\beta), y - x_1 \rangle + \langle \lambda(T(x)) - \lambda(\beta), y - x_1 \rangle \\
&\leq \langle \lambda(T(x)) - \lambda(\beta), y - x_1 \rangle \leq K_1 \|\lambda(T(x))\| \, |T(x) - \beta| \, \|y - x_1\| \\
&\leq K_1 \|\lambda(T(x))\| \, |T(x) - \beta| \big(\|y - x\| + \|x_1 - x\| \big) \\
&\leq K_1 \|\zeta\| \, |T(x) - \beta| \, \big(\|y - x\| + K_2 |T(x) - \beta| \big) \\
&\leq K_3 \|\zeta\| \big(\|y - x\|^2 + |T(x) - \beta|^2 \big)
\end{aligned}$$

for another suitable constant K_3.

Second, we estimate $\langle \zeta, x_1 - x \rangle$. We have

$$\begin{aligned}
\langle \zeta, x_1 - x \rangle &= \int_\beta^{T(x)} \big\langle \lambda(T(x)), F(\bar{x}(s)) + G(\bar{x}(s))\bar{u}(s) \big\rangle \, ds \\
&= \int_\beta^{T(x)} \big\langle \lambda(s), F(\bar{x}(s)) + G(\bar{x}(s))\bar{u}(s) \big\rangle \, ds \\
&+ \int_\beta^{T(x)} \big\langle \lambda(T(x)) - \lambda(s), F(\bar{x}(s)) + G(\bar{x}(x))\bar{u}(s) \big\rangle \, ds \\
&\leq (T(x) - \beta) h(x, \zeta) + K_4 \|\zeta\| |T(x) - \beta|^2,
\end{aligned}$$

for a suitable constant K_4, owing to Corollary 6.4 (applied to the reversed dynamics $\dot{x} = -F(x) - G(x)u$). Therefore, since $h(x\zeta) \leq 0$,

$$\langle (\zeta, \theta), (y, \beta) - (x, T(x)) \rangle \leq (K_3 + K_4) \|\zeta\| \big(\|y - x\|^2 + |T(x) - \beta|^2 \big),$$

and the proof for the case (a) is concluded by observing that K_3 and K_4 are independent of ζ and x.

In the second case we need to use the optimality of x. We observe first that, since $\theta \leq 0$, we only need to prove (6.4) for $\beta = T(y)$. Recalling Corollary 6.3, we can extend the control $\bar{u}$ up to the time $T(y)$ so that the associated trajectory (still denoted by $\bar{x}(\cdot)$) remains optimal. Let also λ be the extended adjoint vector and denote by $\tilde{x}(\cdot)$ the trajectory of the reversed dynamics associated with the extended control $\bar{u}$, i.e.,

$$\begin{cases} \dot{\tilde{x}}(t) = -F(\tilde{x}(t)) - G(\tilde{x}(t))\tilde{u}(t), \\ \tilde{x}(0) = 0, \end{cases}$$

where $\tilde{u}(t) = \bar{u}(T(y) - t)$.

Set $x_1 = \tilde{x}(T(y))$. We estimate first $\langle \zeta, y - x_1 \rangle$. We have, by arguing similarly as before,

$$\begin{aligned}
\langle \zeta, y - x_1 \rangle &= \langle \lambda(T(y)), y - x_1 \rangle + \langle \lambda(T(x)) - \lambda(T(y)), y - x_1 \rangle \\
&\quad \text{(the first summand is } \leq 0 \text{ by the construction in Theorem 6.2)} \\
&\leq \langle \lambda(T(x)) - \lambda(T(y)), y - x_1 \rangle \\
&\leq K_5 \|\zeta\| \big(|T(y) - T(x)|^2 + \|y - x_1\|^2 \big).
\end{aligned}$$

On the other hand,

$$\begin{aligned}\langle \zeta, x_1 - x\rangle &= \int_{T(x)}^{T(y)} \big\langle \zeta, -F(\tilde{x}(s)) - G(\tilde{x}(s))\tilde{u}(s)\big\rangle ds \\ &= \int_{T(x)}^{T(y)} \big\langle \lambda(s), -F(\tilde{x}(s)) - G(\tilde{x}(s))\tilde{u}(s)\big\rangle ds \\ &+ \int_{T(x)}^{T(y)} \big\langle \lambda(T(x)) - \lambda(s), -F(\tilde{x}(s)) - G(\tilde{x}(s))\tilde{u}(s)\big\rangle ds \\ &\leq \int_{T(x)}^{T(y)} \max_{u\in\mathcal{U}} \big\langle \lambda(s), -F(\tilde{x}(s)) - G(\tilde{x}(s))u\big\rangle ds + K_6\|\zeta\|(T(y)-T(x))^2 \\ &\quad \text{(for a suitable constant } K_6 \text{ given by Gronwall's Lemma).}\end{aligned}$$

Recalling Corollary 6.4, the maximized Hamiltonian in the integral of the first summand is constant. Therefore we obtain

$$\langle \zeta, x_1 - x\rangle \leq -h(x,\zeta)(T(y)-T(x)) + K_6\|\zeta\|\,|T(y)-T(x)|^2.$$

Combining the above estimates we obtain finally

$$\langle (\zeta,\theta), (y,T(y)) - (x,T(x))\rangle \leq (K_5+K_6)\|\zeta\|\big(\|y-x\|^2 + |T(y)-T(x)|^2\big),$$

and the proof of the claim is concluded, by observing, again, that K_5, K_6 are independent of x and ζ.

In order to conclude the proof we observe that $N^P_{\mathrm{epi}(T)}$ is pointed at every point $(x,T(x))$, $x\in\mathcal{R}^\tau$, since it is easy to see that the projection of every $(\zeta,\theta)\in N^P_{\mathrm{epi}(T)}(x,T(x))$ onto $\mathbb{R}^N$ is normal to the strictly convex set $\mathcal{R}^\tau$. Therefore, we can apply Corollary 3.1 in [17], with $\Omega_P = \mathrm{int}\mathcal{R}^\tau$, which shows that $\mathrm{epi}(T)$ has positive reach. □

7. Appendix

This section is devoted to some technical lemmas which are used in the proof of the main results.

Lemma 7.1. *Let $K:(a,b)\to[0,1]$ be measurable and let $k\in\mathbb{N}$. Then*

$$\int_a^b (t-a)^k K(t)dt \geq \frac{1}{k+1}\Big(\int_a^b K(t)dt\Big)^{k+1},$$

and

$$\int_a^b (b-t)^k K(t)dt \geq \frac{1}{k+1}\Big(\int_a^b K(t)dt\Big)^{k+1}.$$

Proof. Indeed,

$$\int_a^b (t-a)^k K(t)dt = k!\int_a^b\int_{t_k}^b\dots\int_{t_1}^b K(t_0)dt_0\dots dt_k.$$

Since $K(t)\in[0,1]$ for a.e. $t\in[0,1]$, we obtain that

$$\int_a^b (t-a)^k K(t)dt \geq k!\int_a^b K(t_k)\int_{t_k}^b K(t_{k-1})\dots\int_{t_1}^b K(t_0)dt_0\dots dt_k$$

One can easily prove that

$$\int_a^b K(t_k)\int_{t_k}^b K(t_{k-1})\dots\int_{t_1}^b K(t_0)dt_0\dots dt_k = \frac{1}{(k+1)!}\Big(\int_a^b K(t)dt\Big)^{k+1}$$

The proof is complete. □

Lemma 7.2. *Let $f : [a, b] \to \mathbb{R}$ be absolutely continuous and monotone. Fix $k \geq 0$ and assume that there exists $C \in \mathbb{R}^+$ such that*

$$|f'(s)| \geq C(s-a)^k \quad \forall s \in [a, b]. \tag{7.1}$$

Then, either f has no zeros in (a, b) and then, for all $s \in (a, b)$ either

$$|f(s)| \geq \frac{C}{k+1}(s-a)^{k+1} \quad \text{or} \quad |f(s)| \geq \frac{C}{k+1}(b-s)^{k+1},$$

or there exists $c \in (a, b)$ such that $f(c) = 0$ and then, for all $s \in [a, b]$,

$$|f(s)| \geq \frac{C}{k+1}|c-s|^{k+1}$$

The same conclusions hold if (7.1) is substituted by

$$|f'(s)| \geq C(b-s)^k \quad \forall s \in [a, b].$$

Proof. We treat the case where f is nondecreasing, while the other one can be handled by taking $-f$. So (7.1) now reads as

$$f'(s) \geq C(s-a)^k \quad \forall s \in [a, b].$$

If f has no zeros, we have two cases, namely $f(s) > 0$ for all $s \in (a, b)$ or $f(s) < 0$ for all $s \in (a, b)$. For the first case

$$f(s) - f(a) = \int_a^s f'(t)dt \geq C \int_a^s (t-a)^k dt = \frac{C}{k+1}(s-a)^{k+1},$$

so that

$$f(s) \geq \frac{C}{k+1}(s-a)^{k+1}.$$

In the second case,

$$\begin{aligned} f(b) - f(s) &= \int_s^b f'(t)dt \geq C \int_s^b (t-a)^k dt \\ &= \frac{C}{k+1}\left[(b-a)^{k+1} - (s-a)^{k+1}\right] \geq \frac{C}{k+1}(b-s)^{k+1}. \end{aligned}$$

Therefore,

$$f(s) \leq f(b) - \frac{C}{k+1}(b-s)^{k+1} \leq -\frac{C}{k+1}(b-s)^{k+1}.$$

Assume now that there exists $c \in (a, b)$ such that $f(c) = 0$. Then, for all $s \in [a, c]$ we have

$$-f(s) = f(c) - f(s) = \int_s^c f'(t)dt \geq \frac{C}{k+1}(c-s)^{k+1},$$

while for all $s \in [c, b]$ we have

$$\begin{aligned} f(s) &= f(s) - f(c) = \int_c^s f'(t)dt \geq \frac{C}{k+1}\left[(s-a)^{k+1} - (c-a)^{k+1}\right] \\ &\geq \frac{C}{k+1}(s-c)^{k+1}, \end{aligned}$$

and the proof is complete. □

Proposition 7.3. *Let $K \subset \mathbb{R}^N$ be compact and assume that there exist $\gamma > 0$ and $p > 1$ with the following property: for every $x \in \operatorname{bdry} K$, there exists $\zeta \neq 0$ such that for every $y \in K$ one has*

$$\langle \zeta, y - x \rangle \leq -\gamma \|\zeta\| \|y - x\|^p. \tag{7.2}$$

Then K is convex (with nonempty interior) and, for each $x \in \operatorname{bdry} K$, (7.2) is satisfied by all $\zeta \in N_K(x)$.

Proof. We show first that K is strictly convex. To this aim, assume by contradiction that there exist points $x_1 \neq x_2 \in K$ such that the segment $]x_1, x_2[$ is not contained in the interior of K. Let $0 < t < 1$ be such that $x_t = (1-t)x_1 + tx_2 \in \operatorname{bdry} K$ and let $\zeta \neq 0$ be such that (7.2) holds with x_t in place of x. Obviously, $\langle \zeta, x_2 - x_1 \rangle = 0$, and this is a contradiction. Since K is convex with nonempty interior, for each $x \in \operatorname{bdry} K$ the normal cone $N_K(x)$ is pointed and so it is the convex hull of its exposed rays (see [18]).

We now use Theorem 4.6 in [11] and see that for every unit vector w belonging to an exposed ray of $N_K(x)$, there exists a sequence $x_n \to x$ such that

$$N_K(x_n) = \mathbb{R}^+ w_n, \ \|w_n\| = 1 \quad \text{and } w_n \to w.$$

Of course (7.2) holds with x_n (resp., w_n) in place of x (resp., ζ), so that by passing to the limit, w also satisfies (7.2). By taking convex combinations, we conclude the proof. □

References

[1] M. Bardi & I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*, Birkhäuser, Boston (1997).

[2] U. Boscain & B. Piccoli, *Optimal syntheses for control systems on 2-D manifolds*, Springer, Berlin (2004).

[3] P. Cannarsa, & H. Frankowska, Interior sphere property of attainable sets and time optimal control problems, ESAIM Control Optim. Cal. Var. 12 (2006), 350-370.

[4] P. Cannarsa, F. Marino & P. R. Wolenski, *On the minimum time function for differential inclusions*, Discrete Contin. Dyn. Syst. Ser. A, in print.

[5] P. Cannarsa & Khai T. Nguyen, *Exterior sphere condition and time optimal control for differential inclusions*, SIAM J. Control Optim., to appear.

[6] P. Cannarsa & C. Sinestrari, *Convexity properties of the minimum time function*, Calc. Var. Partial Differential Equations 3 (1995), 273-298.

[7] P. Cannarsa & C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and optimal Control*, Birkhäuser, Boston (2004).

[8] L. Cesari, *Optimization–Theory and Applications*, Springer, New York (1983).

[9] F. H. Clarke *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia (1990).

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

[11] G. Colombo & A. Marigonda, *Differentiability properties for a class of non-convex functions*, Calc. Var. Partial Differential Equations 25 (2006), 1–31.

[12] G. Colombo, A. Marigonda & P. R. Wolenski, *Some new regularity properties for the minimal time function*, SIAM J. Control Optim. 44 (2006), 2285–2299.

[13] G. Colombo & Khai T. Nguyen, *On the structure of the Minimum Time Function*, SIAM J. Control Optim. 48 (2010), 4776–4814.

[14] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. 93 (1959), 418–491.

[15] H. Hermes & J. P. LaSalle, *Functional analysis and time optimal control*, Academic Press, New York-London (1969).

[16] A. Krener, *A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems*, SIAM J. Control 12 (1974), 43–52.

[17] Khai T. Nguyen, *Hypographs satisfying an external sphere condition and the regularity of the minimum time function*, J. Math. Anal. Appl. 372 (2010), 611–628.

[18] R. T. Rockafellar, *Clarke's tangent cones and the boundaries of closed sets in* $\mathbb{R}^n$, Nonlinear Analysis, Theory, Methods and Applications. 3 (1979), 145-154.

[19] H. Sussmann, *A bang-bang theorem with bounds on the number of switchings*, SIAM J. Control Optim. 17 (1979), 629–651.

(Giovanni Colombo) Università di Padova, Dipartimento di Matematica Pura ed Applicata, via Trieste 63, 35121 Padova, Italy

E-mail address: colombo@math.unipd.it

(Khai T. Nguyen) Università di Padova, Dipartimento di Matematica Pura ed Applicata, via Trieste 63, 35121 Padova, Italy

E-mail address: khai@math.unipd.it