# THE MOUNTAIN PASS THEOREM IN TERMS OF TANGENCIES 

SĨ TIỆP ĐINH ${ }^{\dagger}$ AND TIẾN-SƠN PHẠM ${ }^{\ddagger}$


#### Abstract

This paper addresses the Mountain Pass Theorem for locally Lipschitz functions on finite-dimensional vector spaces in terms of tangencies. Namely, let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function with a mountain pass geometry. Let $$
c:=\inf _{\gamma \in \mathcal{A}} \max _{t \in[0,1]} f(\gamma(t))
$$ where $\mathcal{A}$ is the set of all continuous paths joining $x^{*}$ to $y^{*}$. We show that either $c$ is a critical value of $f$ or $c$ is a tangency value at infinity of $f$. This reduces to the Mountain Pass Theorem of Ambrosetti and Rabinowitz in the case where the function $f$ is definable (such as, semi-algebraic) in an o-minimal structure.


## 1. Introduction

The celebrated Mountain Pass Theorem of Ambrosetti and Rabinowitz [2] is a very useful tool in nonlinear analysis with many important applications. For more details, we refer the reader to the comprehensive monographs $[1,16,18,26,28,31]$ with the references therein.

The aim of this paper is to provide a version of the Mountain Pass Theorem for locally Lipschitz functions on finite-dimensional vector spaces in terms of tangencies. To be more precise, let us recall some basic terminology.

Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a $C^{1}$ function and let $x^{*}, y^{*} \in \mathbb{R}^{n}$ be such that there exists an open neighborhood $\mathcal{U}$ of $x^{*}$ satisfying the conditions $y^{*} \notin \overline{\mathcal{U}}$ and

$$
\max \left\{f\left(x^{*}\right), f\left(y^{*}\right)\right\}<\inf _{x \in \mathscr{\mathcal { U }}} f(x)
$$

Consider the family $\mathcal{A}$ of all continuous paths joining $x^{*}$ to $y^{*}$ and set

$$
c:=\inf _{\gamma \in \mathcal{A}} \max _{t \in[0,1]} f(\gamma(t))
$$

The following relation is well-known and has many interesting applications:

$$
c \in K_{0}(f) \cup K_{\infty}(f)
$$

Date: May 15, 2021.
2020 Mathematics Subject Classification. Primary 49J35; Secondary 58C20, 58K05, 58K30.
Key words and phrases. Mountain pass; tangencies; locally Lipschitz; critical values; Clarke subdifferential.
where $K_{0}(f)$ is the set of critical values of $f$ and $K_{\infty}(f)$ is the set of values at which $f$ does not satisfy the weak Palais-Smale condition, i.e.,

$$
K_{0}(f):=\left\{t \in \mathbb{R}: \text { there is a point } x \in \mathbb{R}^{n} \text { such that } \nabla f(x)=0 \text { and } f(x)=t\right\}
$$

and

$$
K_{\infty}(f):=\left\{\begin{array}{cl}
t \in \mathbb{R}: & \text { there is a sequence } x^{k} \rightarrow \infty \text { such that } \\
& f\left(x^{k}\right) \rightarrow t \text { and }\left\|x^{k}\right\|\left\|\nabla f\left(x^{k}\right)\right\| \rightarrow 0
\end{array}\right\}
$$

In a difference line of development, suppose that the function $f$ is polynomial (or more general, definable in an o-minimal structure; see [30] for more on the subject). It is well known that there exists a (minimal) finite set $B(f) \subset \mathbb{R}$, called the bifurcation set of $f$, such that the restriction map

$$
f: \mathbb{R}^{n} \backslash f^{-1}(B(f)) \rightarrow \mathbb{R} \backslash B(f)
$$

is a locally trivial $C^{\infty}$-fibration (see, for example, $[3,11,13,20,25,29,27]$ ). Since $f$ may not be proper, the bifurcation set $B(f)$ contains not only the set of critical values $K_{0}(f)$, but also the set $B_{\infty}(f)$ of atypical values at infinity corresponding to the critical points at infinity. While the set $K_{0}(f)$ is relatively well understood, the other set $B_{\infty}(f)$ is still mysterious. To control the set $B_{\infty}(f)$, we can use the set $T_{\infty}(f)$ of tangency values (at infinity) of $f$ :

$$
T_{\infty}(f):=\left\{\begin{array}{cc}
t \in \mathbb{R}: & \text { there is a sequence } x^{k} \rightarrow \infty \text { such that } \\
& f\left(x^{k}\right) \rightarrow t \text { and } \operatorname{rank}\left\{\nabla f\left(x^{k}\right), x^{k}\right\}=1
\end{array}\right\} .
$$

It is well-known (see, for example, $[10,12,15,17,19]$ ) that $T_{\infty}(f)$ is a finite set and

$$
B_{\infty}(f) \subset T_{\infty}(f) \subset K_{\infty}(f)
$$

The inclusions may be strict (see [21, 22]).

Motivated by the aforementioned works and the usefulness of tangencies in semi-algebraic optimization (see [12, 14, 15, 17, 24] for more details), we will show in Theorem 3.1 that

$$
c \in K_{0}(f) \cup T_{\infty}(f)
$$

Actually, the same conclusion holds even when the function $f$ is not assumed to be differentiable, but merely locally Lipschitz continuous.

The rest of the paper is organized as follows. Section 2 covers some preliminary materials. Section 3 presents the main result and its proof.

## 2. Preliminaries

2.1. Notation. Throughout this work we shall consider the Euclidean vector space $\mathbb{R}^{n}$ endowed with its canonical scalar product $\langle\cdot, \cdot\rangle$, and we shall denote its associated norm $\|\cdot\|$. The open ball (resp., the sphere) centered at $\bar{x} \in \mathbb{R}^{n}$ of radius $r$ will be denoted by $\mathbb{B}_{r}(\bar{x})$ (resp., $\mathbb{S}_{r}(\bar{x})$ ). For simplicity, we write $\mathbb{B}_{r}^{n}$ and $\mathbb{S}_{r}^{n-1}$ if $x=0$; and write $\mathbb{B}^{n}$ and $\mathbb{S}^{n-1}$ if $x=0$ and $r=1$. For a subset $A$ of $\mathbb{R}^{n}$, the closure, the boundary and the convex hull of $A$ are denoted by $\bar{A}, \partial A$ and $\operatorname{co}(A)$ respectively. Let $\operatorname{dist}(A, B)$ stand for the Euclidean distance between $A$ and $B \subset \mathbb{R}^{n}$, namely

$$
\operatorname{dist}(A, B):=\inf \{\|x-y\|: x \in A, y \in B\}
$$

For convenience, if $B \neq \emptyset$, set $\operatorname{dist}(\emptyset, B):=+\infty$.
2.2. Subdifferential of locally Lipschitz mappings. Here we recall the notions and some elementary properties of the Clarke subdifferential and the generalized directional derivative of locally Lipschitz functions used in this paper. The reader is referred to $[6,7,8]$ for more details.

Definition 2.1. Let $F: \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$ be a locally Lipschitz mapping. The Clarke subdifferential of $F$ at $x \in \mathbb{R}^{n}$ is defined by

$$
\partial F(x):=\operatorname{co}\left\{\lim d_{x^{k}} F: x^{k} \rightarrow x \text { and } F \text { is differentiable at } x^{k}\right\},
$$

where $d_{x^{k}} F$ is the differential of $F$ at $x^{k}$, which can be identified with the Jacobian matrix of $F$ at $x^{k}$.

Definition 2.2. Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function and $v \in \mathbb{R}^{n}$. The generalized directional derivative of $f$ at $x$ in the direction $v$, denoted by $f^{\circ}(x ; v)$, is defined as follows:

$$
f^{\circ}(x ; v):=\limsup _{y \rightarrow x, h \rightarrow 0^{+}} \frac{f(y+h v)-f(y)}{h} .
$$

Lemma 2.1. [8, Propositions 2.1.2 and 2.1.5] Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function. Then we have:
(i) For all $x \in \mathbb{R}^{n}$, the set $\partial f(x)$ is a non-empty, convex, compact subset of $\mathbb{R}^{n}$.
(ii) The set-valued mapping $\partial f$ is upper semi-continuous on $\mathbb{R}^{n}$, i.e., for any $x \in \mathbb{R}^{n}$, if $x^{k} \in \mathbb{R}^{n}$ and $w^{k} \in \partial f\left(x^{k}\right)$ are sequences such that $x^{k} \rightarrow x$ and $w^{k} \rightarrow w$, then $w \in \partial f(x)$.
(iii) $f^{\circ}(x ; v)=\max _{w \in \partial f(x)}\langle w, v\rangle$ for any $v \in \mathbb{R}^{n}$.

The following lemma is a slightly changed version of [5, Lemma 3.3].
Lemma 2.2. Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function and $b>0$. Suppose that $U \subset \mathbb{R}^{n}$ is an open set such that

$$
\inf _{w \in \partial f(x)}\|w\| \geqslant 2 b \text { for all } x \in U
$$

Then there exists a locally Lipschitz vector field $\mathrm{v}(x)$ defined on $U$ satisfying

$$
\|\mathrm{v}(x)\|<1 \text { and }\langle w, \mathrm{v}(x)\rangle>b \text { for any } w \in \partial f(x)
$$

Proof. Note that the assumption of the lemma is the conclusion of [5, Lemma 3.2] which is used to prove [5, Lemma 3.3], so the proof is completely similar to that of [5, Lemma 3.3].

In the sequel, we will need the following lemma.
Lemma 2.3. Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function and $D \subset \mathbb{R}^{n}$ be a compact set. For each $\epsilon>0$, there is $h_{0}=h_{0}(\epsilon)>0$ such that for all $y \in N_{h_{0}}(D), h \in\left(0, h_{0}\right]$ and $v \in \overline{\mathbb{B}}^{n}$, we have

$$
\frac{f(y+h v)-f(y)}{h}<f^{\circ}(x ; v)+\epsilon,
$$

where $N_{h_{0}}(D)$ is the closed neighborhood of radius $h_{0}$ of $D$ and $x \in D$ is a point such that $\operatorname{dist}(y, D)=\|y-x\|$.

Proof. By contradiction, assume that there is $\epsilon_{0}>0$ such that for all integer $k>0$, there are $y^{k} \in N_{\frac{1}{k}}(D), x^{k} \in D, h_{k} \in\left(0, \frac{1}{k}\right]$ and $v^{k} \in \overline{\mathbb{B}}^{n}$ such that

$$
\begin{equation*}
\operatorname{dist}\left(y^{k}, D\right)=\left\|y^{k}-x^{k}\right\| \quad \text { and } \quad \frac{f\left(y^{k}+h_{k} v^{k}\right)-f\left(y^{k}\right)}{h_{k}} \geqslant f^{\circ}\left(x^{k} ; v^{k}\right)+\epsilon \tag{1}
\end{equation*}
$$

By taking subsequences if necessary, we can suppose that the sequences $x^{k}$ and $v^{k}$ converge to the limits $x^{0}$ and $v^{0}$, respectively. Since $f$ is locally Lipschitz, there is a neighborhood $U$ of $x^{0}$ and a constant $K>0$ such that $f$ is Lipschitz on $U$ with the constant $K$. In addition, by definition and by shrinking $U$ if necessary, we can suppose that there is $\widetilde{h}>0$ such that

$$
\frac{f\left(y+h v^{0}\right)-f(y)}{h}<f^{\circ}\left(x^{0} ; v^{0}\right)+\frac{\epsilon}{4}
$$

for all $y \in U$ and $h \in(0, \widetilde{h}]$. Set

$$
z^{k}:=y^{k}+h_{k}\left(v^{k}-v^{0}\right) .
$$

It is clear that $y^{k} \rightarrow x^{0}, z^{k} \rightarrow x^{0}$ and $z^{k}+h_{k} v^{0} \rightarrow x^{0}$ as $k \rightarrow+\infty$. Consequently, for $k$ large enough so that $y^{k}, z^{k} \in U, K\left\|v^{k}-v^{0}\right\|<\frac{\epsilon}{4}$ and $h_{k}<\widetilde{h}$, we have

$$
\begin{aligned}
\frac{f\left(y^{k}+h_{k} v^{k}\right)-f\left(y^{k}\right)}{h_{k}}= & \frac{f\left(y^{k}+h_{k}\left(v^{k}-v^{0}\right)+h_{k} v^{0}\right)-f\left(y^{k}+h_{k}\left(v^{k}-v^{0}\right)\right)}{h_{k}} \\
& +\frac{f\left(y^{k}+h_{k}\left(v^{k}-v^{0}\right)\right)-f\left(y^{k}\right)}{h_{k}} \\
= & \frac{f\left(z^{k}+h_{k} v^{0}\right)-f\left(z^{k}\right)}{h_{k}}+\frac{f\left(z^{k}\right)-f\left(y^{k}\right)}{h_{k}} \\
< & f^{\circ}\left(x^{0} ; v^{0}\right)+\frac{\epsilon}{4}+K\left\|v^{k}-v^{0}\right\|<f^{\circ}\left(x^{0} ; v^{0}\right)+\frac{\epsilon}{2} .
\end{aligned}
$$

This contradicts (1) and so ends the proof of the lemma.

## 3. The main result and its proof

For a locally Lipschitz function $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$, we define the set of critical values of $f$ and the set of tangency values (at infinity) of $f$, respectively, by

$$
K_{0}(f):=\left\{t \in \mathbb{R}: \text { there is } x \in f^{-1}(t) \text { such that } 0 \in \partial f(x)\right\}
$$

and

$$
T_{\infty}(f):=\left\{\begin{array}{cl}
t \in \mathbb{R}: & \text { there are sequences } x^{k} \rightarrow \infty \text { and } v^{k} \in \partial f\left(x^{k}\right) \text { such that } \\
& f\left(x^{k}\right) \rightarrow t \text { and } \operatorname{rank}\left\{x^{k}, v^{k}\right\}=1
\end{array}\right\} .
$$

The main result of the paper is as follows.
Theorem 3.1 (Mountain pass). Let $f: \mathbb{R}^{n} \rightarrow \mathbb{R}$ be a locally Lipschitz function and let $x^{*}, y^{*} \in \mathbb{R}^{n}$ with $x^{*} \neq y^{*}$. Assume that there is an open neighborhood $\mathcal{U}$ of $x^{*}$ such that $y^{*} \notin \overline{\mathcal{U}}$ and

$$
f\left(x^{*}\right), f\left(y^{*}\right)<\inf _{x \in \partial \mathcal{U}} f(x)
$$

Let

$$
\begin{equation*}
c:=\inf _{\gamma \in \mathcal{A}} \max _{t \in[0,1]} f(\gamma(t)), \tag{2}
\end{equation*}
$$

where $\mathcal{A}$ stands for the set of all continuous path joining $x^{*}$ to $y^{*}$, i.e.,

$$
\begin{equation*}
\mathcal{A}:=\left\{\gamma \in C\left([0,1], \mathbb{R}^{n}\right): \gamma(0)=x^{*}, \gamma(1)=y^{*}\right\} . \tag{3}
\end{equation*}
$$

Then

$$
c \in K_{0}(f) \cup T_{\infty}(f)
$$

Let us start with some lemmas of preparation.
Lemma 3.1. Let $X \subset \mathbb{R}^{n}$ be a compact set and let $\mathcal{Z}:=\left\{Z_{i}: i=1, \ldots, p\right\}$ be a distinct finite open cover of $X$. Then there exists a constant $\lambda>0$ depending on $\mathcal{Z}$ such that the following statements hold:
(i) For any $i \in\{1, \ldots, p\}$ and any $x \in Z_{i} \cap X$ such that $Z_{i}$ is the unique open set in the cover containing $x$, then $\operatorname{dist}\left(x, \partial Z_{i}\right)>3 \lambda$.
(ii) For any $i \in\{1, \ldots, p\}$ and any $x \in Z_{i} \cap X$ such that $\operatorname{dist}\left(x, \partial Z_{i}\right) \leqslant 3 \lambda$, there is $j \in\{1, \ldots, p\} \backslash\{i\}$ depending on $x$ such that

$$
x \in Z_{j} \quad \text { and } \quad \operatorname{dist}\left(x, \partial Z_{j}\right)>3 \lambda
$$

(iii) Let $i \in\{1, \ldots, p\}$ and $x \in Z_{i} \cap X$ be such that $\operatorname{dist}\left(x, \partial Z_{i}\right)=2 \lambda$ and $\operatorname{dist}(x, X) \leqslant \lambda$. Then for each $y \in X$ with $\|x-y\| \leqslant \lambda$, we have $\operatorname{dist}\left(y, \partial Z_{i}\right) \geqslant \lambda$ and there exists $j \in\{1, \ldots, p\} \backslash\{i\}$ such that

$$
x, y \in Z_{j}, \operatorname{dist}\left(x, \partial Z_{j}\right)>2 \lambda \text { and } \operatorname{dist}\left(y, \partial Z_{j}\right)>3 \lambda .
$$

Proof. (i) By contradiction, assume that for any integer $k>0$, there is an index $i_{k} \in$ $\{1, \ldots, p\}$ and a point $x^{k} \in Z_{i_{k}} \cap X$ such that $Z_{i_{k}}$ is the unique open set in $\mathcal{Z}$ containing $x^{k}$ and $\operatorname{dist}\left(x^{k}, \partial Z_{i_{k}}\right) \leqslant \frac{3}{k}$. Since $\mathcal{Z}$ is finite, by taking a subsequence if necessary, we can suppose that $i_{k}$ is fixed for all $k$, namely, $i_{k}=i \in\{1, \ldots, p\}$. By the compactness of $X$, we can assume that the sequence $x^{k}$ converges to a limit $x^{0} \in X$. Clearly $x^{0} \in \partial Z_{i} \cap X$, in particular, $x^{0} \notin Z_{i}$. Furthermore, $x^{0} \notin Z_{j}$ for all $j \neq i$ since otherwise, $x^{k} \in Z_{j}$ for $k$ large enough which contradicts the fact that $Z_{i}$ is the unique open set in $\mathcal{Z}$ containing $x^{k}$. Consequently, $x^{0} \notin \cup_{j=1, \ldots, p} Z_{j}$. So $x^{0} \notin X$, which is a contradiction.
(ii) Let $\lambda$ be given by item (i). We will show that the statement holds by shrinking $\lambda$. For contradiction, assume that for any integer $k>0$ such that $\frac{1}{k}<\lambda$, there is an index $i_{k} \in\{1, \ldots, p\}$ and a point $x^{k} \in Z_{i_{k}} \cap X$ such that

- $\operatorname{dist}\left(x^{k}, \partial Z_{i_{k}}\right) \leqslant \frac{3}{k}$; and
- for each $j \in\{1, \ldots, p\} \backslash\{i\}$, either $x^{k} \notin Z_{j}$ or $x^{k} \in Z_{j}$ with $\operatorname{dist}\left(x^{k}, \partial Z_{j}\right) \leqslant \frac{3}{k}$.

Since the cover is finite, by taking a subsequence if necessary, we may suppose that $i_{k}$ is fixed for all $k$, namely, $i_{k}=i \in\{1, \ldots, p\}$; moreover, for each $j \in\{1, \ldots, p\} \backslash\{i\}$, either $x^{k} \notin Z_{j}$ for all $k$ or $x^{k} \in Z_{j}$ and $\operatorname{dist}\left(x^{k}, \partial Z_{j}\right) \leqslant \frac{3}{k}$ for all $k$. As $X$ is compact, we may assume that the sequence $x^{k}$ converges to a limit $x^{0} \in X$. Clearly $x^{0} \in \partial Z_{i} \cap X$. For each $j \in\{1, \ldots, p\} \backslash\{i\}$, by construction, either $x^{0} \notin \bar{Z}_{j}$ or $x^{0} \in \bar{Z}_{j}$ with $\operatorname{dist}\left(x^{0}, \partial Z_{j}\right)=0$, in particular, $x^{0} \notin Z_{j}$. Therefore $x^{0} \notin \cup_{j=1, \ldots, p} Z_{j}$. Consequently $x^{0} \notin X$, which is a contradiction.
(iii) Let $\lambda$ be given by (i) and (ii). By assumption, $\mathbb{B}_{2 \lambda}^{n}(x) \subset Z_{i}$. Let $y \in X$ be such that $\|x-y\| \leqslant \lambda$. Then $y \in Z_{i}$. On the other hand,

$$
\operatorname{dist}\left(y, \partial Z_{i}\right) \leqslant\|x-y\|+\operatorname{dist}\left(x, \partial Z_{i}\right) \leqslant 3 \lambda .
$$

By item (ii), there is $j \in\{1, \ldots, p\} \backslash\{i\}$ such that $y \in Z_{j}$ and

$$
\operatorname{dist}\left(y, \partial Z_{j}\right)>3 \lambda
$$

Therefore $\mathbb{B}_{3 \lambda}^{n}(y) \subset Z_{j}$ and so $x \in Z_{j}$. In addition,

$$
\operatorname{dist}\left(x, \partial Z_{j}\right) \geqslant \operatorname{dist}\left(y, \partial Z_{j}\right)-\|x-y\|>2 \lambda
$$

This ends the proof of the lemma.
Lemma 3.2. Let $X \subset \mathbb{R}^{n}$ be a compact set and let $\mathcal{Z}:=\left\{Z_{i}: i=1, \ldots, p\right\}$ be a distinct finite family of open balls covering $X$. Assume that, for each $i$, there is a constant $\rho_{i}>0$, an open set $W_{i} \supset Z_{i}$ and a bi-Lipschitz homeomorphism $\eta_{i}: W_{i} \rightarrow \eta_{i}\left(W_{i}\right) \subset \mathbb{R}^{n}$ such that

$$
\eta_{i}\left(W_{i}\right)=\left\{u \in \mathbb{R}^{n}:\left|u_{j}\right|<\rho_{i}, j=1, \ldots, n\right\} \text { and } \eta_{i}\left(W_{i} \cap X\right) \subset\left\{u \in \mathbb{R}^{n}: u_{2}=0\right\}
$$

Let $\lambda>0$ be the constant depending on $\mathcal{Z}$ given by Lemma 3.1 and $L \geqslant 1$ be a common Lipschitz constant of $\eta_{1}, \ldots, \eta_{p}, \eta_{1}^{-1}, \ldots, \eta_{p}^{-1}$. Let $\gamma:[a, b] \rightarrow \mathbb{R}^{n}$ be a continuous piecewise linear curve such that $\operatorname{dist}(\gamma(t), X) \leqslant \frac{\lambda}{2 L^{2}}$ for $t \in[a, b]$. Then there exist finite sequences
$a=: T_{0}<\cdots<T_{q}:=b$ and $i_{0} \neq i_{1} \neq \cdots \neq i_{q-1}$ with $i_{k} \in\{1, \ldots, p\}(k=0, \ldots, q-1)$ such that:
(i) $\gamma\left[T_{k-1}, T_{k}\right] \subset W_{i_{k-1}}$ for $k=1, \ldots, q$; and
(ii) $\eta_{i_{k}}^{-1}\left(w^{k}\right) \in W_{i_{k-1}} \cap W_{i_{k}}$ for $k=1, \ldots, q-1$, where

$$
z^{k}:=\eta_{i_{k}}\left(\gamma\left(T_{k}\right)\right) \text { and } w^{k}:=\left(z_{1}^{k}, w_{2}^{k}, z_{3}^{k}, \ldots, z_{n}^{k}\right)
$$

with

$$
w_{2}^{k}=\left\{\begin{array}{lll}
-z_{2}^{k} & \text { if } & z_{2}^{k} \neq 0 \\
-\frac{\lambda}{L} & \text { if } & z_{2}^{k}=0
\end{array}\right.
$$

Proof. The construction of the desired sequence is done by induction as follow.
Step 1: $k=0$. Let $y^{0} \in X$ be such that $\left\|y^{0}-\gamma\left(T_{0}\right)\right\|=\operatorname{dist}\left(\gamma\left(T_{0}\right), X\right)<\lambda$. In view of Lemma 3.1(i)-(ii), there is $i_{0} \in\{1, \ldots, p\}$ such that $y^{0} \in Z_{i_{0}}$ and $\operatorname{dist}\left(y^{0}, \partial Z_{i_{0}}\right)>3 \lambda$. So

$$
\operatorname{dist}\left(\gamma\left(T_{0}\right), \partial Z_{i_{0}}\right) \geqslant \operatorname{dist}\left(y^{0}, \partial Z_{i_{0}}\right)-\left\|y^{0}-\gamma\left(T_{0}\right)\right\|>2 \lambda
$$

Hence $\gamma\left(T_{0}\right) \in Z_{i_{0}}$. Set

$$
S_{1}:=\sup \left\{t \in\left[T_{0}, b\right]: \gamma(s) \in Z_{i_{0}} \text { for all } s \in\left[T_{0}, t\right]\right\}>T_{0} .
$$

If $S_{1}=b$, then set $T_{1}:=b$ and we are done. Otherwise, we have $\gamma\left(S_{1}\right) \in \partial Z_{i_{0}}$. So if we let

$$
T_{1}:=\sup \left\{t \in\left[T_{0}, b\right]: \operatorname{dist}\left(\gamma(s), \partial Z_{i_{0}}\right) \geqslant 2 \lambda \text { for all } s \in\left[T_{0}, t\right]\right\},
$$

then clearly

$$
T_{0}<T_{1}<b, \quad \gamma\left[T_{0}, T_{1}\right] \subset W_{i_{0}}, \quad \text { and } \operatorname{dist}\left(\gamma\left(T_{1}\right), \partial Z_{i_{0}}\right)=2 \lambda
$$

By Lemma 3.1(iii), there is $i_{1} \neq i_{0}$ such that $\gamma\left(T_{1}\right) \in Z_{i_{1}}$ and $\operatorname{dist}\left(\gamma\left(T_{1}\right), \partial Z_{i_{1}}\right)>2 \lambda$.
Step 2: Induction. For $k \geqslant 0$, assume that we have constructed sequences $a=$ : $T_{0}<\cdots<T_{k}$, $i_{0} \neq i_{1} \neq \cdots \neq i_{k}$ and $\left\{y^{0}, \ldots, y^{k}\right\}$ such that, for $l=1, \ldots, k$, we have:
(a) $T_{l}:=\sup \left\{t \in\left[T_{l-1}, b\right]: \operatorname{dist}\left(\gamma(s), \partial Z_{i_{l-1}}\right) \geqslant 2 \lambda\right.$ for all $\left.s \in\left[T_{l-1}, t\right]\right\}$;
(b) $\operatorname{dist}\left(\gamma\left(T_{l}\right), \partial Z_{i_{l-1}}\right)=2 \lambda$ and $\operatorname{dist}\left(\gamma\left(T_{l}\right), \partial Z_{i_{l}}\right)>2 \lambda$;
(c) $\gamma\left[T_{l-1}, T_{l}\right] \subset W_{i_{l-1}}$;
(d) $\left\|y^{l-1}-\gamma\left(T_{l-1}\right)\right\|=\operatorname{dist}\left(\gamma\left(T_{l-1}\right), X\right)$ and $\left\|y^{k}-\gamma\left(T_{k}\right)\right\|=\operatorname{dist}\left(\gamma\left(T_{k}\right), X\right)$.

Set

$$
S_{k+1}:=\sup \left\{t \in\left[T_{k}, b\right]: \gamma(s) \in Z_{i_{k}} \text { for all } s \in\left[T_{k}, t\right]\right\}>T_{k}
$$

If $S_{k+1}=1$, set $q:=k+1$ and $T_{k+1}:=b$. Then item (i) follows. Contrarily, we have $\gamma\left(S_{k+1}\right) \in \partial Z_{i_{k}}$. Let

$$
T_{k+1}:=\sup \left\{t \in\left[T_{k}, b\right]: \operatorname{dist}\left(\gamma(s), \partial Z_{i_{k}}\right) \leqslant 2 \lambda \text { for all } s \in\left[T_{k}, t\right]\right\}
$$

then clearly $T_{k}<T_{k+1}<b$ and $\operatorname{dist}\left(\gamma\left(T_{k+1}\right), \partial Z_{i_{k}}\right)=2 \lambda$. By Lemma 3.1(iii), there is $i_{k+1} \neq i_{k}$ such that $\gamma\left(T_{k+1}\right) \in Z_{i_{k+1}}$ and $\operatorname{dist}\left(\gamma\left(T_{k+1}\right), \partial Z_{i_{k+1}}\right)>2 \lambda$. Hence the process can be repeated with $k$ replaced by $k+1$.

Observe that the sequence $a=: T_{0}<T_{1}<\cdots$ is finite so there must be $q>0$ such that $T_{q}=b$. Indeed, suppose for contradiction that the sequence $a=: T_{0}<T_{1}<\cdots$ is infinite. Then there exists an index $k$ that appears infinitely many times in the sequence $i_{0} \neq i_{1} \neq \cdots$. This implies that, in view of item (b), $\gamma$ cuts the sphere $\left\{x \in Z_{k}: \operatorname{dist}\left(z, \partial Z_{k}\right)=2 \lambda\right\}$ infinitely many times which is a contradiction since a piecewise linear curve can meets a sphere at finitely many times. Consequently, (i) follows immediately.

Now we show that $\eta_{i_{k}}^{-1}\left(w^{k}\right) \in W_{i_{k-1}} \cap W_{i_{k}}$ for $k=1, \ldots, q-1$. If $z_{2}^{k}=0$, since $\eta_{i_{k}}^{-1}$ is Lipschitz with the constant $L$, then

$$
\begin{equation*}
\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)-\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\| \leqslant L\left\|z^{k}-w^{k}\right\|=L\left\|z_{2}^{k}-w_{2}^{k}\right\|=\lambda \tag{4}
\end{equation*}
$$

Otherwise, by item (d) and by the assumption $\operatorname{dist}(\gamma(t), X) \leqslant \frac{\lambda}{2 L^{2}}$ for $t \in[a, b]$, we have $\left\|y^{k}-\gamma\left(T_{k}\right)\right\|<\frac{\lambda}{2 L^{2}}$. Hence $\left\|\eta_{i_{k}}\left(y^{k}\right)-z^{k}\right\|<\frac{\lambda}{2 L}$, and so

$$
\begin{align*}
\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)-\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\| & \leqslant\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)-\eta_{i_{k}}^{-1}\left(\eta_{i_{k}}\left(y^{k}\right)\right)\right\|+\left\|\eta_{i_{k}}^{-1}\left(\eta_{i_{k}}\left(y^{k}\right)\right)-\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\| \\
& \leqslant\left\|y^{k}-\gamma\left(T_{k}\right)\right\|+L\left\|\eta_{i_{k}}\left(y^{k}\right)-w^{k}\right\| \\
& =\left\|y^{k}-\gamma\left(T_{k}\right)\right\|+L\left\|\eta_{i_{k}}\left(y^{k}\right)-z^{k}\right\|  \tag{5}\\
& \leqslant\left(1+L^{2}\right)\left\|y^{k}-\gamma\left(T_{k}\right)\right\|<\left(1+L^{2}\right) \frac{\lambda}{2 L^{2}} \leqslant \lambda .
\end{align*}
$$

Now (4) and (5), together with the facts $\operatorname{dist}\left(\gamma\left(T_{k}\right), \partial Z_{i_{k-1}}\right)=2 \lambda$, give

$$
\begin{aligned}
\operatorname{dist}\left(\eta_{i_{k}}^{-1}\left(w^{k}\right), \partial Z_{i_{k-1}}\right) & \geqslant \operatorname{dist}\left(\eta_{i_{k}}^{-1}\left(z^{k}\right), \partial Z_{i_{k-1}}\right)-\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)-\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\| \\
& =\operatorname{dist}\left(\gamma\left(T_{k}\right), \partial Z_{i_{k-1}}\right)-\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)-\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\|>\lambda
\end{aligned}
$$

Similarly, (4), (5) and $\operatorname{dist}\left(\gamma\left(T_{k}\right), \partial Z_{i_{k}}\right)>2 \lambda \operatorname{imply} \operatorname{dist}\left(\eta_{i_{k}}^{-1}\left(w^{k}\right), \partial Z_{i_{k}}\right)>\lambda$. Therefore $\eta_{i_{k}}^{-1}\left(w^{k}\right) \in W_{i_{k-1}} \cap W_{i_{k}}$ for $k=1, \ldots, q-1$. So (ii) follows and the lemma is proved.

We need the following variant of the constant rank theorem for locally Lipschitz mappings.

Lemma 3.3. Let $F: \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$ be a locally Lipschitz mapping with $n \geqslant m$. Assume that each element of $\partial F(x)$ has rank $m$ for any $x$ in a neighborhood of $x^{0} \in \mathbb{R}^{n}$. Then there is an open neighborhood $Z$ of $x^{0}$ in $\mathbb{R}^{n}$ and a bi-Lipschitz homeomorphism $\eta: Z \rightarrow \eta(Z) \subset \mathbb{R}^{n}$ such that

$$
F \circ \eta^{-1}\left(u_{1}, \ldots, u_{n}\right)=\left(u_{1}, \ldots, u_{m}\right)+F\left(x^{0}\right)
$$

for all $\left(u_{1}, \ldots, u_{n}\right) \in \eta(Z)$.
Proof. The proof follows directly from the proof of [4, Theorem 3.1].

Proof of Theorem 3.1. Recall that $\mathcal{A}$ is the set given by (3). For $r \geqslant \max \left\{\left\|x^{*}\right\|,\left\|y^{*}\right\|\right\}$ and $\epsilon>0$, set

$$
\begin{equation*}
\mathcal{A}(r, \epsilon):=\left\{\gamma \in \mathcal{A}: \max _{t \in[0,1]} f(\gamma(t))<c+\epsilon \text { and } \max _{t \in[0,1]}\|\gamma(t)\| \leqslant r\right\} . \tag{6}
\end{equation*}
$$

By definition, for each $\epsilon>0$, there exists $r \gg 1$ such that $\mathcal{A}(r, \epsilon)$ is non-empty so the function

$$
\begin{equation*}
(0,+\infty) \rightarrow(0,+\infty), \quad \epsilon \mapsto R(\epsilon):=\inf \{r: \mathcal{A}(r, \epsilon) \neq \emptyset\} \tag{7}
\end{equation*}
$$

is well-defined and moreover, it is decreasing. In particular, there exists the limit

$$
\lim _{\epsilon \rightarrow 0^{+}} R(\epsilon) \in \mathbb{R} \cup\{+\infty\}
$$

Now Theorem 3.1 is a direct consequence of Propositions 3.1 and 3.2 below.
Proposition 3.1. If $\lim _{\epsilon \rightarrow 0^{+}} R(\epsilon)<+\infty$ then $c \in K_{0}(f)$.
We need some preparation before giving the proof of Proposition 3.1. As $R(\epsilon)$ is decreasing, there is a constant $R_{0}>0$ such that for all $\epsilon>0$, we have $R(\epsilon)<R_{0}$ and so $\mathcal{A}\left(R_{0}, \epsilon\right) \neq \emptyset$. For each integer $k>0$, take $\gamma^{k} \in \mathcal{A}\left(R_{0}, \frac{1}{k}\right)$ and let $D$ be the superior Kuratowski limit of the sequence of non-empty compact sets

$$
D_{k}:=\left\{\gamma^{k}(t): t \in[0,1] \text { and } f\left(\gamma^{k}(t)\right) \geqslant c\right\} .
$$

Namely, $x \in D$ if and only if there is a sequence $x^{k_{l}} \in D_{k_{l}}$ such that $x^{k_{l}} \rightarrow x$ as $l \rightarrow+\infty$. It is clear that $D$ is a non-empty compact set and $f(x)=c$ for any $x \in D$. To prove that $c \in K_{0}(f)$, it is enough to show that there is $x \in D$ such that $0 \in \partial f(x)$. Assume for contradiction that $0 \notin \partial f(x)$ for all $x \in D$. By the compactness of the Clarke subdifferential (Lemma 2.1(i)), it is not hard to see that for each $x \in D$, there is a constant $b_{x}>0$ such that

$$
\inf _{w \in \partial f(x)}\|w\|>4 b_{x}
$$

By Lemma 2.1(ii), there exists a bounded open neighborhood $U_{x}$ of $x$ such that

$$
\inf _{w \in \partial f(y)}\|w\|>2 b_{x} \text { for all } y \in U_{x}
$$

By assumption, we have

$$
f\left(x^{*}\right), f\left(y^{*}\right)<\inf _{x \in \mathscr{\mathcal { U }}} f(x) \leqslant c
$$

Thus $x^{*}, y^{*} \notin D$ and so, we can shrink $U_{x}$ so that $x^{*}, y^{*} \notin U_{x}$.
As $D$ is compact and $\left\{U_{x}: x \in D\right\}$ is an open cover of $D$, there exists a finite open cover of $D$ :

$$
\left\{U_{x^{i}}: x^{i} \in D, i=1, \ldots, p\right\}
$$

Let

$$
b:=\min _{i=1, \ldots, p} b_{x^{i}}>0 \text { and } U:=\bigcup_{i=1}^{p} U_{x^{i}} .
$$

In view of Lemma 2.2, there exists a locally Lipschitz vector field $\mathrm{v}(x)$ defined on $U$ such that

$$
\begin{equation*}
\|\mathrm{v}(x)\|<1 \text { and }\langle w, \mathrm{v}(x)\rangle>b \text { for any } w \in \partial f(x) . \tag{8}
\end{equation*}
$$

Let $h_{0}:=h_{0}\left(\frac{b}{4}\right)$ be the constant determined by Lemma 2.3. So, for all $y \in N_{h_{0}}(D)$, $h \in\left(0, h_{0}\right]$ and $v \in \overline{\mathbb{B}}^{n}$, we have

$$
\begin{equation*}
\frac{f(y+h v)-f(y)}{h}<f^{\circ}(x ; v)+\frac{b}{4}, \tag{9}
\end{equation*}
$$

where $x \in D$ is a point such that $\operatorname{dist}(y, D)=\|y-x\|$. Let $V \subset U$ be an open neighborhood of $D$ such that $\bar{V} \subset U \cap N_{h_{0}}(D)$. By a smooth version of Urysohn's lemma [23, Lemma 1.3.2], there is a smooth function $\varphi: \mathbb{R}^{n} \rightarrow[0,1]$ such that:

$$
\begin{equation*}
\varphi\left(\mathbb{R}^{n} \backslash U\right)=0 \text { and } \varphi(\bar{V})=1 \tag{10}
\end{equation*}
$$

Let

$$
\widetilde{\mathrm{v}}(x):=\varphi(x) \mathrm{v}(x),
$$

which is obviously a locally Lipschitz vector field on $\mathbb{R}^{n}$. This, together with the facts that $U$ is bounded and $\operatorname{supp}(\widetilde{\mathrm{v}}) \subset U$, implies that the vector field $\widetilde{\mathrm{v}}$ is Lipschitz on $\mathbb{R}^{n}$ with a constant $K>0$. Moreover, in view of Lemma 2.1(i), (ii), there is a constant $K^{\prime}>0$ such that $\|w\| \leqslant K^{\prime}$ for all $w \in \partial f(x)$ and $x \in \bar{U}$. We need the following lemma.

Lemma 3.4. For any trajectory $\alpha:\left(t_{1}, t_{2}\right) \rightarrow \mathbb{R}^{n}$ of $-\widetilde{v}$, the function $f \circ \alpha$ is decreasing on $\left(t_{1}, t_{2}\right)$. In addition, for all $u \in\left(t_{1}, t_{2}\right)$ such that $\alpha(u) \in V$, we have

$$
\begin{equation*}
f(\alpha(u+h))-f(\alpha(u))<-\frac{b h}{2} \text { for } h \in\left(0, \min \left\{\frac{b}{2 K K^{\prime}}, h_{0}\right\}\right] \tag{11}
\end{equation*}
$$

where $h_{0}:=h_{0}\left(\frac{b}{4}\right)$ is the constant determined by Lemma 2.3.
Proof. Take any $u \in\left(t_{1}, t_{2}\right)$. For all $s \in\left(t_{1}, t_{2}\right)$, we have

$$
\begin{equation*}
\|\alpha(s)-\alpha(u)\|=\left\|\int_{u}^{s}-\widetilde{\mathrm{v}}\left(\alpha\left(s^{\prime}\right)\right) d s^{\prime}\right\| \leqslant\left|\int_{u}^{s}\left\|\widetilde{\mathrm{v}}\left(\alpha\left(s^{\prime}\right)\right)\right\| d s^{\prime}\right| \leqslant\left|\int_{u}^{s} d s^{\prime}\right|=|s-u| . \tag{12}
\end{equation*}
$$

Observe that the first statement is clear if $\alpha(u) \notin \operatorname{supp} \widetilde{\mathrm{v}} \subset U$ so assume that $\alpha(u) \in \operatorname{supp} \widetilde{\mathrm{v}}$. Thus $\varphi(\alpha(u))>0$ and so $\alpha(u) \in U$. Let

$$
\epsilon:=\frac{\varphi(\alpha(u))}{4} b>0 .
$$

For all $h \in\left(0, \frac{\varphi(\alpha(u)) b}{2 K K^{\prime}}\right]$ small enough, we have

$$
\begin{align*}
\frac{f(\alpha(u+h))-f(\alpha(u))}{h}< & f^{\circ}\left[\alpha(u) ; \frac{\alpha(u+h)-\alpha(u)}{h}\right]+\epsilon \\
= & \max _{w \in \partial f(\alpha(u))}\left\langle w, \frac{\alpha(u+h)-\alpha(u)}{h}\right\rangle+\epsilon \\
= & \max _{w \in \partial f(\alpha(u))} \frac{1}{h}\left\langle w, \int_{u}^{u+h}-\widetilde{\mathrm{v}}(\alpha(s)) d s\right\rangle+\epsilon \\
= & -\min _{w \in \partial f(\alpha(u))} \frac{1}{h}\left\langle w, \int_{u}^{u+h} \widetilde{\mathrm{v}}(\alpha(s)) d s\right\rangle+\epsilon \\
\leqslant & -\min _{w \in \partial f(\alpha(u))} \frac{1}{h}\left\langle w, \int_{u}^{u+h} \widetilde{\mathrm{v}}(\alpha(u)) d s\right\rangle+ \\
& \max _{w \in \partial f(\alpha(u))} \frac{1}{h}\left\langle w, \int_{u}^{u+h}(\widetilde{\mathrm{v}}(\alpha(s))-\widetilde{\mathrm{v}}(\alpha(u))) d s\right\rangle+\epsilon \\
\leqslant & -\min _{w \in \partial f(\alpha(u)))}\langle w, \widetilde{\mathrm{v}}(\alpha(u))\rangle+  \tag{13}\\
& \max _{w \in \partial f(\alpha(u))} \frac{\|w\|}{h} \int_{u}^{u+h}\|\widetilde{\mathrm{v}}(\alpha(s))-\widetilde{\mathrm{v}}(\alpha(u))\| d s+\epsilon \\
\leqslant & \min _{w \in \partial f(\alpha(u))}\langle w, \varphi(\alpha(u)) \mathrm{v}(\alpha(u))\rangle+ \\
& \max _{w \in \partial f(\alpha(u))} \frac{\|w\|}{h} \int_{u}^{u+h} K\|\alpha(s)-\alpha(u)\| d s+\epsilon \\
< & -\varphi(\alpha(u)) b+\frac{K K^{\prime}}{h} \int_{u}^{u+h}(s-u) d s+\epsilon \\
= & -\varphi(\alpha(u)) b+\frac{K K^{\prime}}{h} \int_{u}^{u+h} d\left(\frac{s^{2}}{2}-u s\right)+\epsilon \\
= & -\varphi(\alpha(u)) b+\frac{K K^{\prime} h}{2}+\epsilon,
\end{align*}
$$

where the first inequality follows from the definition of generalized directional derivative, the first equality follows from Lemma 2.1(iii) and the last inequality follows from (8) and (12). Therefore

$$
\begin{equation*}
f(\alpha(u+h))-f(\alpha(u))<-\varphi(\alpha(u)) b h+\frac{K K^{\prime} h^{2}}{2}+\epsilon h \leqslant-\frac{\varphi(\alpha(u))}{2} b h<0 \tag{14}
\end{equation*}
$$

This implies that $f \circ \alpha$ is decreasing at $u$.
It remains to prove the second statement. Assume $\alpha(u) \in V$. In view of (10), we have $\varphi(\alpha(u))=1$. Moreover, $\frac{\alpha(u+h)-\alpha(u)}{h} \in \overline{\mathbb{B}}^{n}$ by (12). Thus, in view of (9), for all $h \in\left(0, \min \left\{\frac{b}{2 K K^{\prime}}, h_{0}\right\}\right]$, by replacing the first inequality in (13) by the following one

$$
\frac{f(\alpha(u+h))-f(\alpha(u))}{h}<f^{\circ}\left[x ; \frac{\alpha(u+h)-\alpha(u)}{h}\right]+\epsilon,
$$

where $x \in D$ is a point such that $\operatorname{dist}(\alpha(u), D)=\|\alpha(u)-x\|$, and repeating the computation in (13), we get

$$
\frac{f(\alpha(u+h))-f(\alpha(u))}{h}<-b+\frac{K K^{\prime} h}{2}+\epsilon .
$$

Consequently

$$
f(\alpha(u+h))-f(\alpha(u))<-b h+\frac{K K^{\prime} h^{2}}{2}+\epsilon h \leqslant-\frac{b h}{2} .
$$

Proof of Proposition 3.1. Let

$$
h:=\min \left\{\frac{b}{2 K K^{\prime}}, h_{0}\right\}>0 .
$$

By construction, it is clear that $\sup _{x \in D_{k}} \operatorname{dist}(x, D) \rightarrow 0$ as $k \rightarrow+\infty$. Therefore, for $k$ large enough,

$$
\begin{equation*}
D_{k} \subset V \quad \text { and } \quad \frac{1}{k}<\frac{h b}{2} \tag{15}
\end{equation*}
$$

Let us fix such an integer $k$. For each $t \in[0,1]$, let $\phi(t, s)$ be the (unique) trajectory of $-\widetilde{v}$ with the initial condition $\phi(t, 0)=\gamma^{k}(t)$, i.e.,

$$
\phi(t, s)=\gamma^{k}(t)-\int_{0}^{s} \widetilde{\mathrm{v}}\left(\phi\left(t, s^{\prime}\right)\right) d s^{\prime}
$$

According to [9, Theorem 9.5], the mapping $\phi$ is continuous with respect to $t$. By construction, $x^{*}, y^{*} \notin U$ and $\widetilde{\mathrm{v}}$ vanishes outside of $U$. So

$$
\phi(0, s)=\gamma^{k}(0)=x^{*} \text { and } \phi(1, s)=\gamma^{k}(0)=y^{*} \text { for all } s .
$$

Consequently $\phi(\cdot, h) \in \mathcal{A}$. We will show that $f(\phi(t, h))<c$ for any $t \in[0,1]$ which contradicts (2). Note that for $t \in[0,1]$ such that $\phi(t, 0)=\gamma^{k}(t) \notin D_{k}$, in light of Lemma 3.4, one has

$$
\begin{equation*}
f(\phi(t, h)) \leqslant f(\phi(t, 0))=f\left(\gamma^{k}(t)\right)<c . \tag{16}
\end{equation*}
$$

Thus it is enough to consider $t \in[0,1]$ such that $\gamma^{k}(t) \in D_{k}$. Observe that for all such $t$, we have $\varphi(\phi(t, 0))=1$ in light of (10) and (15). Therefore, by Lemma 3.4, the fact $\gamma^{k} \in \mathcal{A}\left(R_{0}, \frac{1}{k}\right)$ and (15), we get

$$
\begin{equation*}
f(\phi(t, h))<f(\phi(t, 0))-\frac{b h}{2}<c+\frac{1}{k}-\frac{b h}{2}<c \tag{17}
\end{equation*}
$$

Now, from (16) and (17), it follows that $f(\phi(t, h))<c$ for all $t \in[0,1]$. This contradicts the definition of $c$ given by (2) and so ends the proof of Proposition 3.1.

Proposition 3.2. If $\lim _{\epsilon \rightarrow 0^{+}} R(\epsilon)=+\infty$ then $c \in T_{\infty}(f)$.
Let us make some preparation before proving Proposition 3.2. Recall that the set $\mathcal{A}(r, \epsilon)$ and the real number $R(\epsilon)$ are defined respectively by (6) and (7). Let $\epsilon^{\prime}>0$. For any $\gamma \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)$, we have $\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n} \neq \emptyset$, so the set

$$
I=I(\gamma):=\{t \in[0,1]:\|\gamma(t)\| \geqslant R(\epsilon)\}
$$

is non-empty.

Lemma 3.5. For all $\epsilon>0$ small enough, we have

$$
f^{-1}(c+\epsilon) \neq \emptyset \text { and } R(\epsilon)>\max \left\{\left\|x^{*}\right\|,\left\|y^{*}\right\|\right\} .
$$

Moreover, for all $\epsilon^{\prime}>0$ (depending on $\epsilon$ ) small enough, there is a piecewise linear curve $\gamma \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)$ such that $f(\gamma(t))>c$ for all $t \in I=I(\gamma)$.

Proof. The first statement is clear so let us prove the second one. For this, let $\epsilon^{\prime}>0$ be such that

$$
\begin{equation*}
2 \epsilon^{\prime}<\min \left\{\operatorname{dist}\left(\{f \leqslant c\} \cap \overline{\mathbb{B}}_{R(\epsilon)+1}^{n}, f^{-1}(c+\epsilon / 2)\right), R(\epsilon)-\left\|x^{*}\right\|, R(\epsilon)-\left\|y^{*}\right\|, 2\right\} . \tag{18}
\end{equation*}
$$

Pick an arbitrary $\beta \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)$, we will deform $\beta$ to get the desired curve. Set

$$
g(t):= \begin{cases}\|\beta(t)\| & \begin{array}{l}
\text { if }\|\beta(t)\| \leqslant R(\epsilon)-\epsilon^{\prime} \\
\text { or } f(\beta(t)) \geqslant c+\epsilon / 2
\end{array} \\
\max \left\{\begin{array}{l}
R(\epsilon)-\epsilon^{\prime}, \\
\|\beta(t)\|-\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)
\end{array}\right\} & \text { otherwise. }\end{cases}
$$

Let us show that $g$ is continuous on $[0,1]$. Observe that the function

$$
[0,1] \rightarrow \mathbb{R}, t \mapsto \max \left\{R(\epsilon)-\epsilon^{\prime},\|\beta(t)\|-\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)\right\}
$$

is continuous. Thus, it is clear that $g$ is continuous at any $t \in[0,1]$ such that

$$
\|\beta(t)\| \neq R(\epsilon)-\epsilon^{\prime} \text { and } f(\beta(t)) \neq c+\epsilon / 2
$$

It remains to show that $g$ is continuous at any $t \in[0,1]$ such that

$$
\|\beta(t)\|=R(\epsilon)-\epsilon^{\prime} \quad \text { or } \quad f(\beta(t))=c+\epsilon / 2
$$

Firstly, let $t \in[0,1]$ be such that $\|\beta(t)\|=R(\epsilon)-\epsilon^{\prime}$. Suppose that $t_{k} \in[0,1]$ is a sequence such that $t_{k} \rightarrow t$ with $\left\|\beta\left(t_{k}\right)\right\|>R(\epsilon)-\epsilon^{\prime}$, we need to show that $g\left(t_{k}\right) \rightarrow g(t)$ (note that if such a sequence does not exist, then $g$ is continuous at $t$ obviously). For this, it is enough to assume that

$$
\begin{equation*}
\left\|\beta\left(t_{k}\right)\right\|-\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right)>R(\epsilon)-\epsilon^{\prime} \tag{19}
\end{equation*}
$$

for all $k$ and show that

$$
\left\|\beta\left(t_{k}\right)\right\|-\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right) \searrow R(\epsilon)-\epsilon^{\prime}=\|\beta(t)\|
$$

This is equivalent to show that $\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right) \rightarrow 0$, i.e.,

$$
\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)=0
$$

On the other hand, from (19), we get

$$
\left\|\beta\left(t_{k}\right)\right\|>R(\epsilon)-\epsilon^{\prime}+\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right)
$$

Letting $k \rightarrow+\infty$, we get

$$
\|\beta(t)\| \geqslant R(\epsilon)-\epsilon^{\prime}+\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)=\|\beta(t)\|+\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right) .
$$

Hence dist $\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)=0$ and so $g$ is continuous at any $t \in[0,1]$ such that $\|\beta(t)\|=$ $R(\epsilon)-\epsilon^{\prime}$. Now we show that $g$ is continuous at any $t \in[0,1]$ such that $f(\beta(t))=c+\epsilon / 2$. Let $t_{k} \in[0,1]$ be a sequence such that $t_{k} \rightarrow t$. Without loss of generality, assume that $\|\beta(t)\|>R(\epsilon)-\epsilon^{\prime}$ and $\left\|\beta\left(t_{k}\right)\right\| \neq\|\beta(t)\|$ for all $k$. Note that

$$
\left\|\beta\left(t_{k}\right)\right\|-\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right) \rightarrow\|\beta(t)\|-\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)=\|\beta(t)\| .
$$

Thus, for $k$ large enough,

$$
\left\|\beta\left(t_{k}\right)\right\|-\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right)>R(\epsilon)-\epsilon^{\prime}
$$

Hence, by definition,

$$
g\left(t_{k}\right)=\left\|\beta\left(t_{k}\right)\right\|-\operatorname{dist}\left(\beta\left(t_{k}\right), f^{-1}(c+\epsilon / 2)\right)
$$

which yields $g\left(t_{k}\right) \rightarrow\|\beta(t)\|$. Consequently $g$ is continuous at any $t \in[0,1]$ such that $f(\beta(t))=c+\epsilon / 2$ and so is continuous on $[0,1]$.

Set

$$
\zeta(t):=\left\{\begin{array}{lll}
g(t) \frac{\beta(t)}{\|\beta(t)\|} & \text { if } & \beta(t) \neq 0 \\
0 & \text { if } & \beta(t)=0
\end{array}\right.
$$

We will show that $\zeta$ has the desired properties except being piecewise linear. It is clear that $\zeta(t)$ is continuous and

$$
\|\zeta(t)\| \leqslant\|\beta(t)\| \leqslant R(\epsilon)+\epsilon^{\prime} \text { for any } t \in[0,1]
$$

By (18) and by the definition of the function $g$, it follows that $\zeta(0)=\beta(0)=x^{*}$ and $\zeta(1)=$ $\beta(1)=y^{*}$. Moreover, for all $t \in[0,1]$ such that $\|\beta(t)\|>R(\epsilon)-\epsilon^{\prime}$ and $f(\beta(t))<c+\epsilon / 2$, we have

$$
\begin{aligned}
\|\beta(t)-\zeta(t)\|=\|\beta(t)\|-g(t) & =\min \left\{\|\beta(t)\|-\left(R(\epsilon)-\epsilon^{\prime}\right), \operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)\right\} \\
& \leqslant \operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)
\end{aligned}
$$

This, together with the fact $f(\beta(t))<c+\epsilon / 2$, implies $f(\zeta(t)) \leqslant c+\epsilon / 2$. Consequently $\zeta \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)$. Now we show that $f(\zeta(t))>c$ for all $t \in I(\zeta)$. Assume that $\|\zeta(t)\| \geqslant R(\epsilon)$, then $\|\beta(t)\| \geqslant R(\epsilon)$ and so

$$
\epsilon^{\prime}>\|\beta(t)-\zeta(t)\|=\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right)
$$

Combining this with (18) gives

$$
\begin{aligned}
\operatorname{dist}\left(\zeta(t), f^{-1}(c+\epsilon / 2)\right) & \leqslant\|\beta(t)-\zeta(t)\|+\operatorname{dist}\left(\beta(t), f^{-1}(c+\epsilon / 2)\right) \\
& <2 \epsilon^{\prime}<\operatorname{dist}\left(\{f \leqslant c\} \cap \overline{\mathbb{B}}_{R(\epsilon)+1}^{n}, f^{-1}(c+\epsilon / 2)\right) \\
& \leqslant \operatorname{dist}\left(\{f \leqslant c\} \cap \overline{\mathbb{B}}_{R(\epsilon)+\epsilon^{\prime}}^{n}, f^{-1}(c+\epsilon / 2)\right)
\end{aligned}
$$

which yields $f(\zeta(t))>c$.
Finally, we need to deform $\zeta$ to get the desired curve. Set

$$
\lambda:=\min \left\{c+\epsilon-\max _{t \in[0,1]} f(\zeta(t)), \min _{t \in I(\zeta)} f(\zeta(t))-c\right\}>0
$$

Since $\zeta$ and $f$ are continuous on $[0,1]$ and $\overline{\mathbb{B}}_{R(\epsilon)+\epsilon^{\prime}}^{n}$, respectively, they are uniformly continuous on the respective sets by the Heine-Cantor theorem. Thus there are constants $\nu, \nu^{\prime}>0$ such that

$$
\|\zeta(t)-\zeta(s)\|<\nu \text { for } t, s \in[a, b] \text { with }|t-s|<\nu^{\prime}
$$

and

$$
|f(x)-f(y)|<\lambda \text { for } x, y \in \overline{\mathbb{B}}_{R(\epsilon)+\epsilon^{\prime}}^{n} \text { with }\|x-y\|<\nu
$$

Let $a=: T_{0}<T_{1}<\cdots<T_{q}:=b$ be a finite sequence such that $\left|T_{i}-T_{i-1}\right|<\nu^{\prime}$, for $i=1, \ldots, q$ and let $\gamma:[0,1] \rightarrow \overline{\mathbb{B}}_{R(\epsilon)+\epsilon^{\prime}}^{n}$ be the continuous piecewise linear curve defined by the sequence $\left\{\zeta\left(T_{0}\right), \ldots, \zeta\left(T_{q}\right)\right\}$ so $\gamma(0)=\zeta(0), \gamma(1)=\zeta(1)$. It is not hard to check the following facts:
(a) $\|\gamma(t)\| \leqslant R(\epsilon)+\epsilon^{\prime}$ for any $t \in[0,1]$;
(b) $\max \left\{\left|f(\gamma(t))-f\left(\zeta\left(T_{i-1}\right)\right)\right|,\left|f(\gamma(t))-f\left(\zeta\left(T_{i}\right)\right)\right|\right\}<\lambda$ for $t \in\left[T_{i-1}, T_{i}\right](i=1, \ldots, q)$.

For all $t \in[0,1]$, let $i$ be such that $t \in\left[T_{i-1}, T_{i}\right]$. By (b), we have

$$
\begin{aligned}
f(\gamma(t)) & \leqslant f\left(\gamma\left(T_{i}\right)\right)+\left|f(\gamma(t))-f\left(\gamma\left(T_{i}\right)\right)\right|=f\left(\zeta\left(T_{i}\right)\right)+\left|f(\gamma(t))-f\left(\zeta\left(T_{i}\right)\right)\right| \\
& <f\left(\zeta\left(T_{i}\right)\right)+\lambda \leqslant f\left(\zeta\left(T_{i}\right)\right)+c+\epsilon-\max _{s \in[0,1]} f(\zeta(s)) \leqslant c+\epsilon .
\end{aligned}
$$

Combining this with (a) yields

$$
\gamma \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)
$$

We will show that $f(\gamma(t))>c$ for all $t \in I=I(\gamma)$, which ends the proof of the lemma. Pick arbitrarily $t \in I$ and assume that $\left[\zeta\left(T_{i-1}\right), \zeta\left(T_{i}\right)\right]$ is the line segment containing $\gamma(t)$. Since $\|\gamma(t)\| \geqslant R(\epsilon)$, it follows that $\max \left\{\left\|\zeta\left(T_{i-1}\right)\right\|,\left\|\zeta\left(T_{i}\right)\right\|\right\} \geqslant R(\epsilon)$. Without loss of generality, assume that $\left\|\zeta\left(T_{i}\right)\right\| \geqslant R(\epsilon)$, so $\zeta\left(T_{i}\right) \in I(\zeta)$. From this and (b), we get

$$
\begin{aligned}
f(\gamma(t)) & \geqslant f\left(\gamma\left(T_{i}\right)\right)-\left|f(\gamma(t))-f\left(\gamma\left(T_{i}\right)\right)\right|=f\left(\zeta\left(T_{i}\right)\right)-\left|f(\gamma(t))-f\left(\zeta\left(T_{i}\right)\right)\right| \\
& >f\left(\zeta\left(T_{i}\right)\right)-\lambda \geqslant f\left(\zeta\left(T_{i}\right)\right)-\min _{t \in I(\zeta)} f(\zeta(t))-c \geqslant c .
\end{aligned}
$$

The lemma is proved.
Now the proof needs the following key lemma.
Lemma 3.6. Let $\epsilon>0$ be as in Lemma 3.5. Then there are $x \in X:=f^{-1}([c, c+\epsilon]) \cap \mathbb{S}_{R(\epsilon)}^{n-1}$ and $v \in \partial f(x)$ such that $x$ and $v$ are linear dependent.

Proof. The construction in the proof is illustrated in the figure below.


Consider the mapping $F: \mathbb{R}^{n} \rightarrow \mathbb{R}^{2}$ defined by $F(x):=\left(f(x),\|x\|^{2}\right)$. In light of $[8$, Proposition 2.6.2(e)], for $x \neq 0$, we have

$$
\partial F(x) \subset \partial f(x) \times\{2 x\}=\left\{\binom{v}{2 x}: v \in \partial f(x)\right\} .
$$

Assume for contradiction that $x$ and $v$ are linearly independent for all $x \in X$ and all $v \in$ $\partial f(x)$. Then we have $\operatorname{rank}(w)=2$ for all $x \in X$ and all $w \in \partial F(x)$. In view of Lemma 3.3, for each $x \in X$, there exist an open neighborhood $W_{x}$ of $x$ and a bi-Lipschitz homeomorphism $\eta_{x}: W_{x} \rightarrow \eta_{x}\left(W_{x}\right) \subset \mathbb{R}^{n}$ with $\eta_{x}(x)=0$ such that

$$
\begin{equation*}
F \circ \eta_{x}^{-1}\left(u_{1}, u_{2}, \ldots, u_{n}\right)=\left(u_{1}, u_{2}\right)+F(x)=\left(u_{1}, u_{2}\right)+\left(f(x), R(\epsilon)^{2}\right), \tag{20}
\end{equation*}
$$

for all $u=\left(u_{1}, \ldots, u_{n}\right) \in \eta_{x}\left(W_{x}\right)$. For each $x$, shrinking $W_{x}$ if necessary so that $x^{*}, y^{*} \notin W_{x}$. It is clear that there is $\rho_{x}>0$ such that

$$
B_{x}:=\left\{u \in \mathbb{R}^{n}:\left|u_{i}\right|<\rho_{x}, i=1, \ldots, n\right\} \subset \eta_{x}\left(W_{x}\right)
$$

Shrinking $W_{x}$ more if necessary so that $B_{x}=\eta_{x}\left(W_{x}\right)$. By construction, for any $\left(u_{1}, \ldots, u_{n}\right) \in$ $B_{x},(20)$ is equivalent to the following equalities

$$
\begin{equation*}
f\left(\eta_{x}^{-1}\left(u_{1}, u_{2}, \ldots, u_{n}\right)\right)=u_{1}+f(x) \text { and }\left\|\eta_{x}^{-1}\left(u_{1}, u_{2}, \ldots, u_{n}\right)\right\|^{2}=u_{2}+R(\epsilon)^{2} \tag{21}
\end{equation*}
$$

Let $Z_{x} \subset W_{x}$ be an open ball centered at $x$. As $\left\{Z_{x}: x \in X\right\}$ is an open cover of $X$, by compactness, there are distinct points $\left\{x^{1}, \ldots, x^{p}\right\} \in X$ such that

$$
X \subset \bigcup_{i=1}^{p} Z_{x^{i}}
$$

Let $\lambda>0$ be the constant depending on the cover $\left\{Z_{x^{i}}: i=1, \ldots, p\right\}$ given by Lemma 3.1 and $L \geqslant 1$ be a common Lipschitz constant for $\eta_{1}:=\eta_{x^{1}}, \ldots, \eta_{p}:=\eta_{x^{p}}, \eta_{1}^{-1}, \ldots, \eta_{p}^{-1}$. Set $\nu:=\frac{\lambda}{2 L^{2}}$. Let $\epsilon^{\prime}>0$ and $\gamma$ be, respectively, the constant and the piecewise linear curve depending on $\epsilon^{\prime}$ determined by Lemma 3.5. We will show that if $\epsilon^{\prime}>0$ is small enough, then

$$
\begin{equation*}
\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n} \subset N_{\nu}\left(f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1}\right) . \tag{22}
\end{equation*}
$$

Indeed, assume for contradiction that for all $k>0$ large enough, there is $\gamma^{k} \in \mathcal{A}\left(R(\epsilon)+\frac{1}{k}, \epsilon\right)$ and

$$
y^{k} \in \gamma^{k} \backslash\left(\mathbb{B}_{R(\epsilon)}^{n} \cup N_{\nu}\left(f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1}\right)\right) \neq \emptyset
$$

Observe that $R(\epsilon) \leqslant\left\|y^{k}\right\| \leqslant R(\epsilon)+\frac{1}{k}$. So by compactness, the sequence $y^{k}$ has at least a cluster point, say $y^{0}$. Clearly $\left\|y^{0}\right\|=R(\epsilon)$. In addition, as $c<f\left(y^{k}\right) \leqslant c+\epsilon$ in view of Lemma 3.5, one has $c \leqslant f\left(y^{0}\right) \leqslant c+\epsilon$. Hence

$$
y^{0} \in f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1} .
$$

On the other hand, since $y^{k} \notin N_{\nu}\left(f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1}\right)$, it follows that

$$
\operatorname{dist}\left(y^{k}, f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1}\right)>\nu
$$

By letting $k \rightarrow+\infty$, we get

$$
\operatorname{dist}\left(y^{0}, f^{-1}[c, c+\epsilon] \cap \mathbb{S}_{R(\epsilon)}^{n-1}\right) \geqslant \nu
$$

which is a contradiction. Therefore (22) must hold for $\epsilon^{\prime}>0$ sufficiently small. We will deform $\gamma$ to get an other piecewise linear curve $\widetilde{\gamma} \in \mathcal{A}(R(\epsilon), \epsilon)$ such that $\|\widetilde{\gamma}(t)\|<R(\epsilon)$ for any $t \in[0,1]$. As $\gamma$ is piecewise linear, $\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n}$ has finitely many connected components. Without loss of generality, assume that $\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n}$ is connected. In fact, if $\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n}$ is not connected, then it is enough to apply the process below on each connected component of $\gamma \backslash \mathbb{B}_{R(\epsilon)}^{n}$.

Let $[a, b] \subset(0,1)$ be the interval such that $\gamma(t) \geqslant R(\epsilon)$ if and only if $t \in[a, b]$. In view of Lemma 3.2, there exist finite sequences $a=: T_{0}<\cdots<T_{q}:=b$ and $i_{0} \neq i_{1} \neq \cdots \neq i_{q-1}$ with $i_{k} \in\{1, \ldots, p\}(k=0, \ldots, q)$ such that:
(a) $\gamma\left[T_{k-1}, T_{k}\right] \subset W_{i_{k-1}}$ for $k=1, \ldots, q$; and
(b) $\eta_{i_{k}}^{-1}\left(w^{k}\right) \in W_{i_{k-1}} \cap W_{i_{k}}$ for $k=1, \ldots, q-1$, where

$$
z^{k}:=\eta_{i_{k}}\left(\gamma\left(T_{k}\right)\right) \text { and } w^{k}:=\left(z_{1}^{k}, w_{2}^{k}, z_{3}^{k}, \ldots, z_{n}^{k}\right)
$$

with

$$
w_{2}^{k}=\left\{\begin{array}{ccc}
-z_{2}^{k} & \text { if } & z_{2}^{k} \neq 0 \\
-\frac{\lambda}{L} & \text { if } & z_{2}^{k}=0
\end{array}\right.
$$

Since $\gamma$ is continuous, for $\delta>0$ small enough, we have

$$
\begin{gathered}
a-\delta \in[0,1], b+\delta \in[0,1] \\
\gamma(a-\delta) \in W_{i_{0}} \text { and } \gamma(b+\delta) \in W_{i_{q-1}} .
\end{gathered}
$$

Let

$$
\begin{gather*}
a-\delta=: \widetilde{T}_{0}<\widetilde{T}_{1}:=T_{1}<\cdots<\widetilde{T}_{q-1}:=T_{q-1}<\widetilde{T}_{q}:=b+\delta, \\
w^{0}:=\eta_{i_{0}}\left(\gamma\left(\widetilde{T}_{0}\right)\right) \in \eta_{i_{0}}\left(W_{i_{0}}\right)=B_{x_{i_{0}}}, \widetilde{w}^{q}:=\eta_{i_{q-1}}\left(\gamma\left(\widetilde{T}_{q}\right)\right) \in \eta_{i_{q-1}}\left(W_{i_{q-1}}\right)=B_{x_{i_{q-1}}} \tag{23}
\end{gather*}
$$

and

$$
\widetilde{w}^{k}:=\eta_{i_{k-1}}\left(\eta_{i_{k}}^{-1}\left(w^{k}\right)\right) \text { for } k=1, \ldots, q-1 .
$$

Define the curve

$$
\widetilde{\gamma}(t):= \begin{cases}\gamma(t) & \text { if } t \in[0,1] \backslash\left(\widetilde{T}_{0}, \widetilde{T}_{q}\right) \\ \eta_{i_{k-1}}^{-1}\left[w^{k-1}+\frac{t-\widetilde{T}_{k-1}}{\widetilde{T}_{k}-\widetilde{T}_{k-1}}\left(\widetilde{w}^{k}-w^{k-1}\right)\right] & \text { if } t \in\left[\widetilde{T}_{k-1}, \widetilde{T}_{k}\right](k=1, \ldots, q) .\end{cases}
$$

We note the following facts:

- $\widetilde{\gamma}\left(\widetilde{T}_{0}\right)=\eta_{i_{0}}^{-1}\left(w^{0}\right)=\gamma\left(\widetilde{T}_{0}\right)$,
- $\eta_{i_{k}}^{-1}\left(w^{k}\right)=\eta_{i_{k-1}}^{-1}\left(\widetilde{w}^{k}\right)$ for $k=1, \ldots, q-1$,
- $\widetilde{\gamma}\left(\widetilde{T}_{q}\right)=\eta_{i_{q-1}}^{-1}\left(\widetilde{w}^{q}\right)=\gamma\left(\widetilde{T}_{q}\right)$, and
- $w^{k-1}, \widetilde{w}^{k} \in B_{x_{i_{k}}}$ for $k=1, \ldots, q$.

Thus $\widetilde{\gamma}$ is continuous. We will prove that $\|\widetilde{\gamma}(t)\|<R(\epsilon)$ and $f(\widetilde{\gamma}(t))<c+\epsilon$ for any $t \in[0,1]$, which contradicts the definition of $R(\epsilon)$. Clearly,

$$
\|\widetilde{\gamma}(t)\|<R(\epsilon) \quad \text { and } \quad f(\widetilde{\gamma}(t))<c+\epsilon \text { for all } t \in[0,1] \backslash\left(\widetilde{T}_{0}, \widetilde{T}_{q}\right) .
$$

So it remains to show that $\|\widetilde{\gamma}(t)\|<R(\epsilon)$ and $f(\widetilde{\gamma}(t))<c+\epsilon$ for all $t \in\left(\widetilde{T}_{0}, \widetilde{T}_{q}\right)$.
This and (23) implies

$$
\left\|\eta_{i_{0}}^{-1}\left(w^{0}\right)\right\|=\left\|\gamma\left(\widetilde{T}_{0}\right)\right\|<R(\epsilon) \text { and }\left\|\eta_{i_{q-1}}^{-1}\left(\widetilde{w}^{q}\right)\right\|=\left\|\gamma\left(\widetilde{T}_{q}\right)\right\|<R(\epsilon)
$$

On the other hand, by (21),

$$
\left\|\eta_{i_{0}}^{-1}\left(w^{0}\right)\right\|^{2}=w_{2}^{0}+R(\epsilon)^{2} \text { and }\left\|\eta_{i_{q-1}}^{-1}\left(\widetilde{w}^{q}\right)\right\|^{2}=\widetilde{w}_{2}^{q}+R(\epsilon)^{2} .
$$

Hence

$$
\begin{equation*}
w_{2}^{0}<0 \text { and } \widetilde{w}_{2}^{q}<0 \tag{24}
\end{equation*}
$$

In light of (21), (b) and the fact $\widetilde{T}_{k}=T_{k} \in[a, b]=I(\gamma)$ for $k=1, \ldots, q-1$, we have

$$
z_{2}^{k}+R(\epsilon)^{2}=\left\|\eta_{i_{k}}^{-1}\left(z^{k}\right)\right\|^{2}=\left\|\gamma\left(\widetilde{T}_{k}\right)\right\|^{2} \geqslant R(\epsilon)^{2} .
$$

So $z_{2}^{k} \geqslant 0$. By construction, it follows that

$$
\begin{equation*}
w_{2}^{k}<0 \text { for } k=1, \ldots, q-1 \tag{25}
\end{equation*}
$$

Hence

$$
\widetilde{w}_{2}^{k}+R(\epsilon)^{2}=\left\|\eta_{i_{k-1}}^{-1}\left(\widetilde{w}^{k}\right)\right\|^{2}=\left\|\eta_{i_{k}}^{-1}\left(w^{k}\right)\right\|^{2}=w_{2}^{k}+R(\epsilon)^{2}<R(\epsilon)^{2}
$$

where the first and third equalities follows from (21). This yields

$$
\begin{equation*}
\widetilde{w}_{2}^{k}<0 \text { for } k=1, \ldots, q-1 \tag{26}
\end{equation*}
$$

Combining (24), (25) and (26), we get

$$
\begin{equation*}
w_{k-1}, \widetilde{w}^{k} \in B_{x_{i_{k-1}}} \cap\left\{u \in \mathbb{R}^{n}: u_{2}<0\right\} \text { for } k=1, \ldots, q \tag{27}
\end{equation*}
$$

As $\gamma \in \mathcal{A}\left(R(\epsilon)+\epsilon^{\prime}, \epsilon\right)$, we get

$$
f\left(\gamma\left(\widetilde{T}_{0}\right)\right)<c+\epsilon \quad \text { and } \quad f\left(\gamma\left(\widetilde{T}_{q}\right)\right)<c+\epsilon
$$

On the other hand, by $(21)$ and the facts $\gamma\left(\widetilde{T}_{0}\right) \in W_{i_{0}}$ and $\gamma\left(\widetilde{T}_{q}\right) \in W_{i_{q-1}}$ (by (23)), we get

$$
f\left(\gamma\left(\widetilde{T}_{0}\right)\right)=w_{1}^{0}+f\left(x^{i_{0}}\right) \text { and } f\left(\gamma\left(\widetilde{T}_{q}\right)\right)=\widetilde{w}_{1}^{q}+f\left(x^{i_{q-1}}\right)
$$

Thus

$$
\begin{equation*}
w_{1}^{0}<c+\epsilon-f\left(x^{i_{0}}\right) \text { and } \widetilde{w}_{1}^{q}<c+\epsilon-f\left(x^{i_{q-1}}\right) \tag{28}
\end{equation*}
$$

In view of $(21)$, for $k=1, \ldots, q-1$, we have

$$
f\left(\eta_{i_{k}}^{-1}\left(w^{k}\right)\right)=w_{1}^{k}+f\left(x^{i_{k}}\right)=z_{1}^{k}+f\left(x^{i_{k}}\right)=f\left(\eta_{i_{k}}^{-1}\left(z^{k}\right)\right)=f\left(\gamma\left(\widetilde{T}_{k}\right)\right)<c+\epsilon
$$

So

$$
\begin{equation*}
w_{1}^{k}<c+\epsilon-f\left(x^{i_{k}}\right) \text { for } k=1, \ldots, q-1 \tag{29}
\end{equation*}
$$

In addition,

$$
\widetilde{w}_{1}^{k}+f\left(x^{i_{k-1}}\right)=f\left(\eta_{i_{k-1}}^{-1}\left(\widetilde{w}^{k}\right)\right)=f\left(\eta_{i_{k}}^{-1}\left(w^{k}\right)\right)<c+\epsilon
$$

This implies

$$
\begin{equation*}
\widetilde{w}_{1}^{k}<c+\epsilon-f\left(x^{i_{k-1}}\right) \text { for } k=1, \ldots, q-1 \tag{30}
\end{equation*}
$$

From (28), (29) and (30), we get

$$
\begin{equation*}
w_{k-1}, \widetilde{w}^{k} \in B_{x_{i_{k-1}}} \cap\left\{u \in \mathbb{R}^{n}: u_{1}<c+\epsilon-f\left(x^{i_{k-1}}\right)\right\} \text { for } k=1, \ldots, q \tag{31}
\end{equation*}
$$

Now for $k=1, \ldots, q$, by (27) and (31), we get

$$
w^{k-1}, \widetilde{w}^{k} \in B_{x_{i_{k-1}}} \cap\left\{u \in \mathbb{R}^{n}: u_{1}<c+\epsilon-f\left(x^{i_{k-1}}\right), u_{2}<0\right\}
$$

By convexity, $B_{x_{i_{k-1}}} \cap\left\{u \in \mathbb{R}^{n}: u_{1}<c+\epsilon-f\left(x^{i_{k-1}}\right), u_{2}<0\right\}$ also contains the segment $\left[w^{k-1}, \widetilde{w}^{k}\right]$ joining $w^{k-1}$ and $\widetilde{w}^{k}$. Clearly for any $w=\left(w_{1}, \ldots, w_{n}\right) \in\left[w^{k-1}, \widetilde{w}^{k}\right]$, we have

$$
w_{1}<c+\epsilon-f\left(x^{i_{k-1}}\right) \text { and } w_{2}<0
$$

This and (21) yield

$$
f\left(\eta_{i_{k-1}}^{-1}(w)\right)=w_{1}+f\left(x^{i_{k-1}}\right)<c+\epsilon \text { and }\left\|\eta_{i_{k-1}}^{-1}(w)\right\|=\sqrt{w_{2}+R(\epsilon)^{2}}<R(\epsilon)
$$

Consequently $\|\widetilde{\gamma}(t)\|<R(\epsilon)$ for any $t \in\left(\widetilde{T}_{0}, \widetilde{T}_{q}\right)$ and so this holds for all $t \in[0,1]$. This contradiction ends the proof of the lemma.

Now we are in position to finish the proof of Proposition 3.2.
Proof of Proposition 3.2. For each integer $k>0$ large enough such that $\epsilon=\frac{1}{k}$ satisfies the assumptions of Lemma 3.5, in view of Lemma 3.6, there are $x^{k} \in X_{k}:=f^{-1}\left[c, c+\frac{1}{k}\right] \cap \mathbb{S}_{R\left(\frac{1}{k}\right)}^{n-1}$ and $v^{k} \in \partial f\left(x^{k}\right)$ such that $x^{k}$ and $v^{k}$ are linear dependent. Since $R\left(\frac{1}{k}\right) \rightarrow+\infty$ as $k \rightarrow+\infty$, we have $x^{k} \rightarrow \infty$. Furthermore, it is clear that $f\left(x^{k}\right) \rightarrow c$. Consequently $c \in T_{\infty}(f)$.

## References

[1] A. Ambrosetti and A. Malchiodi. Nonlinear Analysis and Semilinear Elliptic Problems. Cambridge University Press, Cambridge, 2007.
[2] A. Ambrosetti and P. H. Rabinowitz. Dual variational methods in critical point theory and applications. J. Functional Analysis, 14:349-381, 1973.
[3] Broughton. Milnor numbers and the topology of polynomial hypersurfaces. Invent. Math., 92:217-241, 1988.
[4] G. J. Butler, J. G. Timourian, and C. Viger. The rank theorem for locally Lipschitz continuous functions. Canad. Math. Bull., 31(2):217-226, 1988.
[5] K. C. Chang. Variational methods for nondifferentiable functionals and their applications to partial differential equations. J. Math. Anal. Appl., 80(1):102-129, 1981.
[6] F. H. Clarke. A new approach to Lagrange multipliers. Math. Oper. Res., 1(2):165-174, 1976.
[7] F. H. Clarke. Generalized gradients of Lipschitz functionals. Adv. in Math., 40(1):52-67, 1981.
[8] F. H. Clarke. Optimization and nonsmooth analysis. Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1990.
[9] R. Coleman. Calculus on normed vector spaces. Springer, New York, 2012.
[10] D. D'Acunto. Valeurs critiques asymptotiques d'une fonction définissable dans une structure o-minimale. Ann. Polon. Math., 75(1):35-45, 2000.
[11] H. V. Hà and D. T. Lê. Sur la topologie des polynômes complexes. Acta Math. Vietnam., 9(1):21-32, 1984.
[12] H. V. Hà and T. S. Phạm. Global optimization of polynomials using the truncated tangency variety and sums of squares. SIAM J. Optim., 19(2), 2008.
[13] H. V. Hà and T. S. Phạm. On the Łojasiewicz exponent at infinity of real polynomials. Ann. Polon. Math., 94(3), 2008.
[14] H. V. Hà and T. S. Phạm. Solving polynomial optimization problems via the truncated tangency variety and sums of squares. J. Pure Appl. Algebra, 213:2167-2176, 2009.
[15] H. V. Hà and T. S. Phạm. Genericity in polynomial optimization, volume 3 of Series on Optimization and Its Applications. World Scientific, Singapore, 2017.
[16] Y. Jabri. The Mountain Pass Theorem. Cambridge Univ. Press, Cambridge, 2003.
[17] D. S. Kim, T. S. Phạm, and N. V. Tuyen. On the existence of Pareto solutions for polynomial vector optimization problems. Math. Program. Ser. A, 177(1-2):321-341, 2019.
[18] 1. Nirenberg. Variational and topological methods in nonlinear problems. Bull. Amer. Math. Soc. (N.S.), $4(3): 267-302,1981$.
[19] T. L. Loi and A. Zaharia. Bifurcation sets of functions definable in o-minimal structures. Illinois J. Math., 43(3):449-457, 1998.
[20] A. Parusiński. On the bifurcation set of complex polynomial with isolated singularities at infinity. Compositio Math., 97(3):369-384, 1995.
[21] L. Păunescu and A. Zaharia. On the Łojasiewicz exponent at infinity for polynomial functions. Kodai Math. J., 20:269-274, 1997.
[22] L. Păunescu and A. Zaharia. Remarks on the Milnor fibration at infinity. manuscripta math., 103:351361, 2000.
[23] P. Petersen. Manifold Theory. https://www.math.ucla.edu/~petersen/manifolds.pdf.
[24] T. S. Phạm. Local minimizers of semi-algebraic functions from the viewpoint of tangencies. SIAM J. Optim., 30(3):1777-1794, 2020.
[25] P. Rabier. Ehresmann fibrations and Palais-Smale conditions for morphisms of Finsler manifolds. Ann. of Math., 146(3):647-691, 1997.
[26] P. Rabinowitz. Minimax Methods in Critical Point Theory with Applications to Differential Equations. CBMS Reg. Conf. Ser. in Math. No. 65, AMS, Providence, 1986.
[27] D. Siersma and M. Tibăr. Singularities at infinity and their vanishing cycles. Duke Math. J., 80(3):771783, 1995.
[28] M. Struwe. Variational Methods. Springer, Berlin, 1996.
[29] R. Thom. Ensembles et morphismes stratifiés. Bull. Amer. Math. Soc., 75:249-312, 1969.
[30] L. van den Dries. Tame topology and o-minimal structures, volume 248 of London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1998.
[31] M. Willem. Minimax Theorems. Birkhäuser, Basel, 1996.

Institute of Mathematics, VAST, 18, Hoang Quoc Viet Road, Cau Giay District 10307, Hanoi, Vietnam

Email address: dstiep@math.ac.vn

Department of Mathematics, Dalat University, 1 Phu Dong Thien Vuong, Dalat, Viet-
NAM
Email address: sonpt@dlu.edu.vn

