

A MORSE-SARD THEOREM FOR THE DISTANCE FUNCTION ON RIEMANNIAN MANIFOLDS

LUDOVIC RIFFORD

ABSTRACT. We prove that the set of critical values of the distance function from a submanifold of a complete Riemannian manifold is of Lebesgue measure zero. In this way, we extend a result of Itoh and Tanaka.

1. INTRODUCTION

A well-known and widely used property of critical values of differentiable mappings is given by the Morse-Sard theorem [11, 15]: if a mapping is C^k -smooth with k sufficiently big, then the set of its critical values has the Lebesgue measure zero. In this article, we prove that the Morse-Sard theorem holds when the smooth function is replaced by the distance function from a C^∞ -smooth submanifold in a complete C^∞ -smooth Riemannian manifold of any dimension. Therefore we settle an open question asked by Itoh and Tanaka [8] who proved this result in dimension less than 5. Our proof is mainly based on the Yomdin's method which relies on semialgebraic sets, and on a result by Itoh and Tanaka [7], which established the Lipschitz continuity of the distance function to the cut locus. Notice that our main theorem can indeed be viewed as a corollary of a more general result about the minimum of smooth functions on compact manifolds. In order to state our results, we present the notion of critical points of the distance function which is not smooth but only Lipschitz continuous.

Let N be a closed C^∞ -smooth submanifold of a complete C^∞ -smooth Riemannian manifold M . The distance function $d_N(\cdot)$ from the submanifold N is defined on M by

$$\forall p \in M, d_N(p) := \min\{d(p, q) : q \in N\}.$$

A unit speed geodesic segment $\gamma : [0, a] \rightarrow M$ emanating from N is called a N -segment if $t = d_N(\gamma(t))$ on $[0, a]$. A point $p \in M \setminus N$ is called a critical point of the distance function d_N if for any unit tangent vector v at p , there exists an N -segment $\gamma : [0, d_N(x)] \rightarrow M$ through p such that the angle made by v and $-\dot{\gamma}(d_N(p))$ is not greater than $\frac{\pi}{2}$; we denote by $C(d_N)$ the set of critical points of the distance function d_N . Notice that this definition of critical point is equivalent to say that the point p contains zero in its Clarke generalized gradient $\partial d_N(p)$; we will return to this characterization of critical

1991 *Mathematics Subject Classification.* 53C22, 28A78, 49J52.

L. Rifford: Institut Girard Desargues, Université Claude Bernard Lyon I, 21 avenue Claude Bernard, 69622 Villeurbanne, France. Email: rifford@desargues.univ-lyon1.fr.

points at the end of this section. In this article, we prove that the set of all critical values of d_N has the Lebesgue measure zero.

Theorem 1. *Let N be a closed C^∞ -smooth submanifold of a complete C^∞ -smooth Riemannian manifold M . Then the set of critical values of the distance function from N is of Lebesgue measure zero.*

Remark 1.1. *Twenty years ago, Fu proved in [6] that the set of all critical values of the distance function from a compact subset in \mathbb{R}^3 is of Lebesgue measure zero. Following Whitney [17], S. Ferry [5] gave a counterexample to this result in dimension greater than three.*

Actually, we are able to improve Theorem 1 whenever the ambient manifold is the Euclidean space \mathbb{R}^n .

Theorem 2. *Let N be a closed submanifold of dimension d and of class C^k in \mathbb{R}^n with $k > 2d(n + 1)$. Then the set of critical values of the distance function from N is of Lebesgue measure zero.*

We do not know if the lower bound on k that we give in Theorem 2 is optimal. Note that both theorems above imply that for almost all positive numbers $c < \sup\{d_N(x)\}$, the set of all points whose distances from N are c is a Lipschitz hypersurface of M . Actually, Theorem 1 is a consequence of a more general result that we proceed to explain.

Let N be a compact smooth manifold and let $\phi : \mathbb{R}^n \times N \rightarrow \mathbb{R}$ be a smooth function. Define the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\forall x \in \mathbb{R}^n, \quad f(x) := \min_{q \in N} \{\phi(x, q)\}.$$

Though the function f is not smooth, we are able to define some notion of critical points for f . Since the manifold N is compact, it is straightforward to show that f is indeed Lipschitz continuous. Thus by Rademacher Theorem (see [12]), it is differentiable almost everywhere in \mathbb{R}^n . Denote by D_f the set of points where f is differentiable and define the generalized gradient of f at $x \in \mathbb{R}^n$ as follows:

$$\forall x \in \mathbb{R}^n, \quad \partial f(x) = \text{co} \left\{ \lim_{i \rightarrow \infty} \nabla f(x_i) : x_i \rightarrow_{i \rightarrow \infty} x, x_i \in D_f \right\},$$

where $\text{co}\{A\}$ denotes the convex hull of $A \subset \mathbb{R}^n$. This tool was introduced by Clarke [2] in 1973. Since that time, a complete calculus has been developed, one that extends all the theorems of the usual smooth calculus. A point $x \in \mathbb{R}^n$ will be called a critical point of f if $0 \in \partial f(x)$, the critical values of f are the images of its critical points, and we denote by $C(f)$ the set of critical points of the function f . Since the function f is defined as the minimum of smooth function, we can compute its generalized gradients (we refer the reader to [3, Exercise 9.13 p. 99] for such a result):

$$\forall x \in \mathbb{R}^n, \quad \partial f(x) = \text{co}\{\nabla_x \phi(x, q) : q \in \text{argmin} f(x)\}, \quad (1)$$

where $\nabla_x \phi(x, q)$ denotes the derivative of ϕ in the first variable and where $\text{argmin} f(x) := \{q \in N : f(x) = \phi(x, q)\}$. In the context of distance functions, as we will see in the proof of Theorem 1, this formula will imply that

both definitions of critical points coincide. In consequence, Theorem 1 will be a corollary of the following.

Theorem 3. *Let \mathcal{U} be an open subsets of \mathbb{R}^n and let N be a compact manifold of class C^k and of dimension d . Let $\phi : \mathcal{U} \times N \rightarrow \mathbb{R}$ be a smooth function of class C^k . Then the function $f : \mathcal{U} \rightarrow \mathbb{R}$ defined by*

$$f(x) := \min_{q \in N} \{\phi(x, q)\},$$

satisfies the following: If $k > 2n + 2d(n + 1)$, then $f(C(f))$ has Lebesgue measure zero.

As we said, our proofs are strongly based on the Yomdin's proof of Sard theorem; they demonstrate the power of the Yomdin's method. Our paper is organized as follows: in Section 2 we develop preliminaries on semialgebraic sets, and we give the proofs of our results in Section 3.

Throughout this paper, \mathbb{R} denotes the set of real numbers, $\|\cdot\|$ the Euclidean norm of \mathbb{R}^n , B^n the open ball $\{x \in \mathbb{R}^n : \|x\| < 1\}$, $\overline{B^n}$ the closure of B and $B^n(x, r) = x + rB^n$ (resp. $\overline{B^n}(x, r) = x + r\overline{B^n}$) the ball (resp. the closed ball) centered at x with radius r . Finally if A is a subset of \mathbb{R}^n then $\text{co}A$ denotes its convex hull.

2. PRELIMINARIES ON SEMIALGEBRAIC SETS

2.1. Definitions and properties. We refer the reader to [4, 10] for surveys on semialgebraic sets and for the proofs of the propositions below. Moreover we refer the interested reader to the excellent book [1] for a very detailed study of semialgebraic sets.

A set $A \subset \mathbb{R}^n$ is called semialgebraic if it can be obtained by a finite number of the union and the intersection operations from the sets of the form $\{f = 0\}$, $\{g > 0\}$, where f, g are polynomials on \mathbb{R}^n . For given such a representation of A we call its set-theoretic formula together with the dimension n of the ambient space and with the degrees of polynomials in it, the diagram of this representation of A . Hence any constant which depends only on the diagram of A can indeed be replaced by another constant which depends only on the number and the degrees of polynomials defining A . We give below three classical propositions on semialgebraic sets; we usually say that these results are consequences of the Tarski-Seidenberg principle.

Proposition 2.1. *The number of connected components of a semialgebraic set is bounded by a constant which depends only on the number and the degrees of polynomials which define that set.*

Proposition 2.2. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set. Then the sets \overline{A} , ∂A , each connected component of A , $\pi(A)$, where π denotes the projection of \mathbb{R}^n on $\mathbb{R}^k \subset \mathbb{R}^n$, are semialgebraic, and the diagram of each of these sets depends only on the diagram of A .*

Proposition 2.3. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set. If P_1, \dots, P_q are q polynomials on \mathbb{R}^n of degree less or equal than d , then the set*

$$\{(P_1(a), \dots, P_q(a)) : a \in A\}$$

is a semialgebraic set in \mathbb{R}^q ; its diagram depends only on the diagram of A , d and q .

Our results will be strongly based on the following theorem; we refer the reader to Yomdin's papers [18] and [19] for the proof of it.

Theorem 4. *For any semialgebraic set $A \subset \mathbb{R}^n$ inside the ball of a given radius r , there is a constant K depending only on the number and degrees of polynomials defining A , such that any x, y belonging to the same connected component of A , can be joined by a connected semialgebraic curve s of length not greater than Kr .*

2.2. The semialgebraic case: quasi critical values. Let S be a positive integer. Let A_1, \dots, A_S be S compact semialgebraic sets in \mathbb{R}^m and let $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$ be S polynomials on $\mathbb{R}^n \times \mathbb{R}^m$. We define $f_P : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\forall x \in \mathbb{R}^n, \quad f_P(x) := \min \{P_s(x, a) : a \in A_s, s \in \{1, \dots, S\}\}.$$

Since the sets A_1, \dots, A_S are compact, the function f_P is locally Lipschitz on \mathbb{R}^n . Let δ, μ be two positive constants, let us define the notion of (δ, μ) -critical points of f_P . For every $x \in \mathbb{R}^n$, we denote by $\operatorname{argmin}^\delta(x)$ the set of couples (a, s) in $\mathbb{R}^m \times \mathbb{R}$ such that

$$s \in \{1, \dots, S\}, a \in A_s, \text{ and } P_s(x, a) \leq f_P(x) + \delta.$$

We will say that $x \in \mathbb{R}^n$ is a (δ, μ) -critical point of f_P if the following property is satisfied:

$$\mu \overline{B} \cap \operatorname{co} \left\{ \nabla_x P_s(x, a) : (a, s) \in \operatorname{argmin}^\delta(x) \right\} \neq \emptyset.$$

We denote by $C_{\delta, \mu}(f_P)$ the set of (δ, μ) -critical points of the function f_P and we prove the following result.

Theorem 5. *Let A_1, \dots, A_S be S compact semialgebraic sets in \mathbb{R}^m and $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$ be S polynomials in $\mathbb{R}^n \times \mathbb{R}^m$. Denote by $K_P > 0$ some uniform Lipschitz constant of the polynomials P_1, \dots, P_S on the set $\overline{B}^n \times (A_1 \cup \dots \cup A_S)$ and denote by $M_P > 0$ some upper bound of the $\|D_x^2 P_1\|, \dots, \|D_x^2 P_S\|$ on the same set¹. Then for any constants δ, μ, r such that $0 < \delta, \mu \leq r \leq 1$, the set $f_P(C_{\delta, \mu}(f_P) \cap \overline{B}_r^n)$ is included in the union of N intervals of length*

$$Kr\mu + \sqrt{2M_P}Kr\delta + K_P \sqrt{\frac{2}{M_P}} \sqrt{\delta},$$

where the constants N and K depend only on the diagrams of the semialgebraic sets A_1, \dots, A_S , on the integer S and of the degrees of the polynomials $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$.

Proof.

Lemma 1. *The set $C_{\delta, \mu}(f_P)$ is semialgebraic; its diagram depends only on the diagrams of A_1, \dots, A_S , on S and on the degrees of polynomials $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$.*

¹Here if $P(\cdot, \cdot)$ is a polynomial on $\mathbb{R}^n \times \mathbb{R}^m$, then $D_x^2 P$ denotes the second order derivative of the polynomial P in the first variable.

Proof. By the Carathéodory theorem (cf. Theorem 17.1 in [13]), the set $C_{\delta, \mu}(f_P)$ can be described as the set of points $x \in \mathbb{R}^n$ which satisfy the two following conditions:

- (i) there exist s_1, \dots, s_{n+1} in $\{1, \dots, S\}$, there exist a_1, \dots, a_{n+1} in \mathbb{R}^m such that,

$$\forall i \in \{1, \dots, n+1\}, \quad a_i \in A_{s_i},$$

$$\text{and } \forall s \in \{1, \dots, S\}, \forall a \in A_s, \quad P_{s_i}(x, a_i) \leq P_s(x, a) + \delta;$$

- (ii) there exists also $(\lambda_1, \dots, \lambda_{n+1})$ in $[0, 1]$ with $\sum_{i=1}^{n+1} \lambda_i = 1$ such that

$$\left\| \sum_{i=1}^{n+1} \lambda_i \nabla_x P_{s_i}(x, a_i) \right\|^2 \leq \mu^2$$

Denote by \mathcal{E} the set of tuples $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1})$ in $\mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1}$ such that

$$\forall i \in \{1, \dots, n+1\}, \quad a_i \in A_{s_i}.$$

Since the sets A_1, \dots, A_S are semialgebraic, the set \mathcal{E} can be represented as a finite union (over all possible $(n+1)$ -tuples (s_1, \dots, s_{n+1})) of semialgebraic sets (of the form $\mathbb{R}^n \times \{s_1\}, \dots, \{s_{n+1}\} \times A_{s_1} \times \dots \times A_{s_{n+1}} \times [0, 1]^{n+1}$), so it is a semialgebraic set. Moreover its diagram depends only on the diagrams of the A_1, \dots, A_S .

Let F be the set of tuples $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1}, s, a)$ in $\mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1} \times \{1, \dots, S\} \times \mathbb{R}^m$ such that

$$\forall i \in \{1, \dots, n+1\}, \quad a_i \in A_{s_i},$$

$$a \in A_s,$$

$$\text{and } \exists i \in \{1, \dots, n+1\} \text{ s.t. } P_{s_i}(x, a_i) > P_s(x, a) + \delta.$$

By the same argument as above, since $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$ are all polynomials, the set F is semialgebraic and its diagram depends only on data of the problem.

Denote by π_1 the projection from $\mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1} \times \{1, \dots, S\} \times \mathbb{R}^m$ into $\mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1}$ which suppress the two last ‘‘coordinates’’ of the tuple, that is, which maps the tuple $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1}, s, a)$ to the tuple $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1})$.

It is straightforward to show that the set $\mathcal{E} \setminus \pi_1(F)$ corresponds to the set of tuples $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1})$ in $\mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1}$ which satisfy

$$\forall i \in \{1, \dots, n+1\}, \quad a_i \in A_{s_i}, \tag{2}$$

and $\forall i \in \{1, \dots, n+1\}, \forall s \in \{1, \dots, S\}, \forall a \in A_s, P_{s_i}(x, a_i) \leq P_s(x, a) + \delta$, which can also be written as,

$$\forall i \in \{1, \dots, n+1\}, \quad P_{s_i}(x, a_i) \leq f_P(x) + \delta. \tag{3}$$

Therefore, if we denote by π_2 the projection

$$\begin{aligned} \pi_2 : \mathbb{R}^n \times \{1, \dots, S\}^{n+1} \times (\mathbb{R}^m)^{n+1} \times [0, 1]^{n+1} &\longrightarrow \mathbb{R}^n \\ (x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1}) &\longmapsto x \end{aligned}$$

then (by (2), (3)) the set $C_{\delta,\mu}(f_P)$ corresponds to $\pi_2(\mathcal{E} \cap \mathcal{E}')$ where the set \mathcal{E}' denotes the set of tuples $(x, s_1, \dots, s_{n+1}, a_1, \dots, a_{n+1}, \lambda_1, \dots, \lambda_{n+1})$ such that

$$\sum_{i=1}^{n+1} \lambda_i = 1,$$

$$\text{and } \left\| \sum_{i=1}^{n+1} \lambda_i \nabla_x P_{s_i}(x, a_i) \right\|^2 \leq \mu^2.$$

Since the projection of a semialgebraic set is semialgebraic by Proposition 2.2, and since all inequalities that appear in (i)-(ii) are polynomial in the coordinates of their variables, we deduce that $C_{\delta,\mu}(f_P)$ is semialgebraic and that its diagram depends only upon that of the sets A_1, \dots, A_S , on the integer S and of the degrees of the polynomials $P_1(\cdot, \cdot), \dots, P_S(\cdot, \cdot)$. \square

Returning to the proof of Theorem 5. Denote by D_i each connected component of $C_{\delta,\mu}(f_P) \cap \overline{B_r^n}$ and let $E_i := f_P(D_i)$. Let x, y be two points belonging to the same connected component of D_i . By Theorem 4 there exist some constant $K > 0$ independent on i and a curve S (parametrized by arc-length) of length $l \leq Kr$, joining x and y in D_i . Set $\rho := \mu + \sqrt{2M_P\delta}$. Let $0 \leq t_i < t_{i+1} \leq l$ such that $t_{i+1} - t_i = \frac{\rho - \mu}{M_P}$; we are going to prove that

$$f_P(x(t_{i+1})) - f_P(x(t_i)) \leq \rho(t_{i+1} - t_i). \quad (4)$$

We argue by contradiction; so let us assume that

$$f_P(x(t_{i+1})) - f_P(x(t_i)) > \rho(t_{i+1} - t_i).$$

Since $x(t_i) \in C_{\delta,\mu}(f_P)$, there exists $(\bar{a}, \bar{s}) \in \operatorname{argmin}^\delta(x(t_i))$ such that

$$\langle x(t_{i+1}) - x(t_i), \nabla_x P_{\bar{s}}(x(t_i), \bar{a}) \rangle \leq \mu \|x(t_{i+1}) - x(t_i)\|.$$

By Taylor formula at second order, we deduce

$$\begin{aligned} P_{\bar{s}}(x(t_{i+1}), \bar{a}) &\leq P_{\bar{s}}(x(t_i), \bar{a}) + \langle \nabla_x P_{\bar{s}}(x(t_i), \bar{a}), x(t_{i+1}) - x(t_i) \rangle \\ &\quad \dots + \frac{M_P}{2} \|x(t_{i+1}) - x(t_i)\|^2, \end{aligned}$$

which, since $P_{\bar{s}}(x(t_i), \bar{a}) \leq f_P(x(t_i)) + \delta$ and $f_P(x(t_{i+1})) \leq P_{\bar{s}}(x(t_{i+1}), \bar{a})$ implies,

$$f_P(x(t_{i+1})) \leq f_P(x(t_i)) + \delta + \mu \|x(t_{i+1}) - x(t_i)\| + \frac{M_P}{2} \|x(t_{i+1}) - x(t_i)\|^2.$$

Remark that since the curve S is parametrized by arc-length, we have $\|x(t_{i+1}) - x(t_i)\| \leq (t_{i+1} - t_i)$. Hence we deduce

$$f_P(x(t_{i+1})) - f_P(x(t_i)) \leq \delta + \mu(t_{i+1} - t_i) + \frac{M_P}{2}(t_{i+1} - t_i)^2.$$

Since we are assuming $f_P(x(t_{i+1})) - f_P(x(t_i)) > \rho(t_{i+1} - t_i)$, this implies

$$(t_{i+1} - t_i) \left[\rho - \mu - \frac{M_P}{2}(t_{i+1} - t_i) \right] < \delta. \quad (5)$$

By definition of ρ and $t_{i+1} - t_i$, the left-hand side of (5) equals δ hence we get a contradiction. We conclude that for every interval $I = [t_i, t_{i+1}] \subset [0, l]$ of length $\frac{\rho - \mu}{M_P}$, we have $f_P(x(t_{i+1})) - f_P(x(t_i)) \leq \rho(t_{i+1} - t_i)$. In particular, this

implies that on each subinterval of $[0, l]$ of the form $I = [0, t]$ with $t = k \frac{\rho - \mu}{M_P}$ and where k is a positive integer, we have

$$f_P(x(t)) - f_P(x(0)) \leq \rho t \leq \rho l \leq \rho K r.$$

The interval $[0, l]$ on which the curve S is defined, can be partitioned into an interval of length of the form $k \frac{\rho - \mu}{M_P}$ and another subinterval of length less than $\frac{\rho - \mu}{M_P}$. Since the polynomial P is Lipschitz with constant K_P on $\overline{B^n} \times (A_1 \cup \dots \cup A_S)$, the function f_P has the same Lipschitz constant on $\overline{B^n}$, hence we deduce that

$$\begin{aligned} f_P(y) - f_P(x) &\leq K r \rho + K_P \frac{\rho - \mu}{M_P} \\ &= K r \mu + \sqrt{2 M_P} K r \delta + K_P \sqrt{\frac{2}{M_P}} \sqrt{\delta}. \end{aligned}$$

Since x and y are arbitrary points in D_i and since the number of connected components of $C_{\delta, \mu}(f_P) \cap \overline{B_r^n}$ depends only on the diagrams of A and $P(\cdot, \cdot)$, we conclude. \square

3. PROOF OF OUR RESULTS

3.1. Proof of Theorem 3. Fix some compact cube \mathcal{T} in \mathcal{U} and let us prove that $\text{meas}(f(C(f) \cap \mathcal{T})) = 0$. Without loss of generality, we can assume that the n -cube \mathcal{T} has sides of length 1.

By Whitney embedding's theorem (see [16, Theorem 4.4 p. 63]), without loss of generality, we can assume that the manifold N is a compact submanifold of \mathbb{R}^m of dimension d and of class C^k . Therefore the function ϕ can be smoothly extended to some neighbourhood \mathcal{V} of N in \mathbb{R}^m in such a way that the function

$$\begin{aligned} \phi : \mathcal{U} \times \mathcal{V} &\longrightarrow \mathbb{R} \\ (x, q) &\longmapsto \phi(x, q) \end{aligned}$$

is a C^k -smooth function. Let $\mathcal{V}' \subset \mathcal{V}$ be a compact neighbourhood of N ; for the rest of the proof we denote by K_ϕ the Lipschitz constant of the function ϕ on $\mathcal{T} \times \mathcal{V}'$ and we denote by R_ϕ^2 (resp. R_ϕ^k) the maximum of the norms $\|d(\nabla_x \phi)\|$ (resp. $\|d^k \phi(x, q)\|$) for $x \in \mathcal{T}$ and $q \in \mathcal{V}'$.

Furthermore, there exists a covering of N by a finite family of open sets $(O_i)_{i \in I}$ in \mathcal{V}' such that for each $i \in I$, O_i is included in some cube of side 1 and such that there exists a diffeomorphism $\Psi_i : \mathcal{C}^m \rightarrow O_i$ of class C^k which satisfies

$$\Psi_i(\mathcal{C}^m \cap (\mathbb{R}^d \times \{0_{m-d}\})) = O_i \cap N,$$

where \mathcal{C}^m is some cube of dimension m in \mathbb{R}^m with sides of length less or equal than one (in fact for sake of simplicity, we can assume that \mathcal{C}^m is the open unit cube of \mathbb{R}^m). From now we denote by K_Ψ the maximum of Lipschitz constants of the mappings Ψ_i ($i \in I$) and by R_Ψ^k the maximum of the $\|d^k \Psi_i\|$ ($i \in I$) over \mathcal{C}^m (without loss of generality, we can assume that the Ψ_i 's are defined on some uniform compact set containing the cube \mathcal{C}^m). Let l be a positive integer, set $r := \frac{1}{l}$. The cube $\mathcal{C}^d := \mathcal{C}^m \cap (\mathbb{R}^d \times \{0_{m-d}\})$ can be divided into l^d cubes \mathcal{D}_j ($j = 1, \dots, l^d$) of side r . For each $i \in I$ and

each j , let $\mathcal{P}_{i,j}$ be the $(k-1)$ -Taylor polynomial² of Ψ_i at the center of the cube \mathcal{D}_j . We have that for any $x \in \mathcal{D}_j$,

$$\|\Psi_i(x) - \mathcal{P}_{i,j}(x)\| \leq R_\Psi^k r^k. \quad (6)$$

By Proposition 2.3, for each i, j , the set $\mathcal{P}_{i,j}(\mathcal{D}_j)$ is semialgebraic in \mathbb{R}^m . Moreover by definition of K_Ψ , there exists $c_{i,j} \in \mathcal{P}_{i,j}(\mathcal{D}_j)$ such that

$$\mathcal{P}_{i,j}(\mathcal{D}_j) \subset \overline{B^n}(c_{i,j}, K_\Psi r). \quad (7)$$

For sake of simplicity, let us set for each couple (i, j) , $\mathcal{E}_{i,j} := \mathcal{P}_{i,j}(\mathcal{D}_j)$. The cube \mathcal{T} in \mathcal{U} can be divided in smaller cubes as well. Thus there exists l^n cubes \mathcal{T}_t ($t = 1, \dots, l^n$) of side r which cover \mathcal{T} ; denote by x_t the center of each cube \mathcal{T}_t . For each (i, j, t) , let $P_{i,j}^t(x, q)$ be the $(k-1)$ -th Taylor polynomial of $\phi(\cdot, \cdot)$ at the point $(x_t, c_{i,j})$. By (7) and by the definition of R_ϕ^k , we have for every $x \in \mathcal{T}_t$ and for every $q \in \mathcal{E}_{i,j}$,

$$|\phi(x, q) - P_{i,j}^t(x, q)| \leq R_\phi^k K_\Psi^k r^k \quad (8)$$

$$\|\nabla_x \phi(x, q) - \nabla_x P_{i,j}^t(x, q)\| \leq R_\phi^k K_\Psi^{k-1} r^{k-1} \quad (9)$$

$$\|d(\nabla_x \phi)(x, q) - d(\nabla_x P_{i,j}^t)(x, q)\| \leq R_\phi^k K_\Psi^{k-2} r^{k-2}. \quad (10)$$

In this way, we produced one family of polynomials $\{P_{i,j}^t\}$ (with $i \in I$, $j = 1, \dots, l^d$ and $t = 1, \dots, l^n$) of degree $k-1$ and one family $\{\mathcal{E}_{i,j}\}$ (with $i \in I$ and $j = 1, \dots, l^d$) of semialgebraic sets which are all defined by polynomials of degree $k-1$. Let us call a configuration, one $(2n+3)$ -tuple of the form $(i_1, i_2, \dots, i_{n+1}, j_1, j_2, \dots, j_{n+1}, t)$ where $t \in \{1, \dots, l^n\}$, where each i_s ($s = 1, \dots, n+1$) belongs to I , and where each j_s ($s = 1, \dots, n+1$) belongs to the set $\{1, \dots, l^d\}$. We denote by \mathcal{S} a configuration which corresponds to some $(2n+3)$ -tuple of the form above. Notice that there are exactly $|I|^{n+1} l^{d(n+1)} l^n$ such $(2n+3)$ -tuples. We claim the following.

Lemma 2. *There exist two positive constants M_1, M_2 such that if the point x belongs to $C(f) \cap \mathcal{T}$, then there exists some configuration \mathcal{S} such that x is a $(M_1 r^k, M_2 r^{k-1})$ -critical point of the mapping*

$$\begin{aligned} f_{\mathcal{S}} : \mathcal{T}_t &\longrightarrow \mathbb{R} \\ x &\longmapsto \min\{P_{i_s, j_s}^t(x, q_s) : q_s \in \mathcal{E}_{i_s, j_s}, s \in \{1, \dots, n+1\}\}. \end{aligned}$$

Moreover we have,

$$|f(x) - f_{\mathcal{S}}(x)| \leq M_1 r^k. \quad (11)$$

Proof. First, since x is in \mathcal{T} , then there exists $t \in \{1, \dots, l^n\}$ such that $x \in \mathcal{T}_t$. Moreover, since it is a critical value of f , by (1) there exist $\bar{q}_1, \dots, \bar{q}_{n+1}$ in the manifold N and $\lambda_1, \dots, \lambda_{n+1} \in [0, 1]^{n+1}$ with $\sum_{s=1}^{n+1} \lambda_s = 1$ such that

$$\sum_{s=1}^{n+1} \lambda_s \nabla_x \phi(x, \bar{q}_s) = 0, \quad (12)$$

$$\text{with } f(x) = \phi(x, \bar{q}_s), \quad \forall s = 1, \dots, n+1. \quad (13)$$

²The function Ψ_i is defined from $\mathcal{C}^m \subset \mathbb{R}^m$ into \mathbb{R}^m , hence its Taylor polynomial denotes the m -tuple with coordinates each Taylor polynomials of the coordinates of Ψ_i .

Thus each \bar{q}_s belongs to some set O_i with $i \in I$. This means that there exist $n+1$ couples $(i_1, j_1), \dots, (i_{n+1}, j_{n+1})$ and $n+1$ points x_s in \mathcal{D}_{j_s} ($s = 1, \dots, n+1$) such that for every $s = 1, \dots, n+1$,

$$\Psi_{i_s}(x_s) = \bar{q}_s.$$

Set for each $s = 1, \dots, n+1$, let $\bar{q}'_s := \mathcal{P}_{i_s, j_s}(x_s)$. Each \bar{q}'_s belongs to \mathcal{E}_{i_s, j_s} and moreover by (6),

$$\forall s = 1, \dots, n+1, \|\bar{q}_s - \bar{q}'_s\| \leq R_\Psi^k r^k. \quad (14)$$

By (9)-(10) and (12) we obtain,

$$\begin{aligned} & \left\| \sum_{s=1}^{n+1} \lambda_s \nabla_x P_{i_s, j_s}^t(x, \bar{q}'_s) \right\| \\ & \leq \left\| \sum_{s=1}^{n+1} \lambda_s \nabla_x P_{i_s, j_s}^t(x, \bar{q}_s) \right\| \\ & \quad + \left\| \sum_{s=1}^{n+1} \lambda_s (\nabla_x P_{i_s, j_s}^t(x, \bar{q}'_s) - \nabla_x P_{i_s, j_s}^t(x, \bar{q}_s)) \right\| \\ & \leq R_\phi^k K_\Psi^{k-1} r^{k-1} + (R_\phi^2 + R_\phi^k K_\Psi^{k-2} r^{k-2}) R_\Psi^k r^k. \end{aligned} \quad (15)$$

Furthermore, by (8), (13) and (14), we have

$$\begin{aligned} f_{\mathcal{S}}(x) - f(x) &= f_{\mathcal{S}}(x) - \phi(x, \bar{q}_1) \\ &\leq f_{\mathcal{S}}(x) - \phi(x, \bar{q}'_1) + K_\phi R_\Psi^k r^k \\ &\leq P_{i_1, j_1}^t(x, \bar{q}'_1) - \phi(x, \bar{q}'_1) + K_\phi R_\Psi^k r^k \\ &\leq R_\phi^k K_\Psi^k r^k + K_\phi R_\Psi^k r^k. \end{aligned} \quad (16)$$

On the other hand, if $q = \mathcal{P}_{i, j}(p)$ belongs to $\mathcal{E}_{i, j}$ then since $\Psi_i(p) \in O_i \cap N$ we get by (8),

$$\begin{aligned} |P_{i, j}^t(x, q) - \phi(x, \Psi_i(p))| &\leq |P_{i, j}^t(x, q) - \phi(x, q)| + K_\phi R_\Psi^k r^k \\ &\leq R_\phi^k K_\Psi^k r^k + K_\phi R_\Psi^k r^k. \end{aligned} \quad (17)$$

Hence we deduce that for some $\bar{q} = \mathcal{P}_{\bar{i}, \bar{j}}(\bar{p}) \in \mathcal{E}_{\bar{i}, \bar{j}}$ such that $f_{\mathcal{S}}(x) = P_{\bar{i}, \bar{j}}^t(x, \bar{q})$, we have

$$\begin{aligned} f(x) - f_{\mathcal{S}}(x) &\leq f(x) - P_{\bar{i}, \bar{j}}^t(x, \bar{q}) \\ &\leq \phi(x, \Psi_{\bar{i}}(\bar{p})) - P_{\bar{i}, \bar{j}}^t(x, \bar{q}) \\ &\leq R_\phi^k K_\Psi^k r^k + K_\phi R_\Psi^k r^k. \end{aligned} \quad (18)$$

Consequently, by (13), (17) and (18), we obtain that for every $s = 1, \dots, n+1$,

$$\begin{aligned} P_{i_s, j_s}^t(x, \bar{q}'_s) &\leq [P_{i_s, j_s}^t(x, \bar{q}_s) - \phi(x, \bar{q}_s)] + \phi(x, \bar{q}_s) \\ &\leq R_\phi^k K_\Psi^k r^k + K_\phi R_\Psi^k r^k + f(x) \\ &\leq f_{\mathcal{S}}(x) + 2(R_\phi^k K_\Psi^k + K_\phi R_\Psi^k) r^k. \end{aligned} \quad (19)$$

Set

$$M_1 := 2(R_\phi^k K_\Psi^k + K_\phi R_\Psi^k),$$

and

$$M_2 := R_\phi^k K_\Psi^{k-1} + (R_\phi^2 + R_{\phi_k} K_\Psi^{k-2} r^{k-2}) R_\Psi^k r.$$

By (15) and (19), we conclude that x is a $(M_1 r^k, M_2 r^{k-1})$ -critical point of f_S . Moreover by (16) and (18), we conclude that

$$|f(x) - f_S(x)| \leq M_1 r^k.$$

The proof of the lemma is complete. \square

Returning to the proof of Theorem 3. Since the semialgebraic sets $\mathcal{E}_{i,j}$ are all defined by a fix number of polynomials of degree $k-1$, since the polynomials $P_{i,j}^t$ have degree $k-1$ and since their Lipschitz constants and their C^2 norms are uniform on $\mathcal{T} \times \mathcal{V}'$ by (8)-(10) (let us call them K_P and M_P); by Theorem 5, we deduce that there exists two positive constants N and K such that for each configuration \mathcal{S} , $f(C(f_S) \cap \mathcal{T}_t)$ is included in N intervals of length

$$Kr\mu + \sqrt{2M_P} Kr\delta + K_P \sqrt{\frac{2}{M_P}} \sqrt{\delta},$$

where $\delta := M_1 r^k$ and $\mu := M_2 r^{k-1}$. By (11) we conclude that for r small enough, we have

$$\begin{aligned} \text{meas}(f(C(f) \cap \mathcal{T})) &\leq |I|^{n+1} l^{d(n+1)} l^n 2K_P \sqrt{\frac{2M_1}{M_P}} \frac{1}{\sqrt{l^k}} \\ &\leq |I|^{n+1} 2K_P \sqrt{\frac{2M_1}{M_P}} l^{n+d(n+1)-\frac{k}{2}}. \end{aligned}$$

As l tends to infinity, since $k > 2n + 2(n+1)d$, we deduce that

$$\text{meas}(f(C(f) \cap \mathcal{T})) = 0,$$

which conclude the proof of Theorem 3.

3.2. Proof of Theorem 1. For each critical point p of the distance function $d_N(\cdot)$ we will show that there exists an open neighbourhood U of p such that the Lebesgue measure of $d_N(U \cap C(d_N))$ is zero. Then $C(d_N)$ is covered by a union of countably many subsets U_i ($i \in \mathbb{N}$) such that the Lebesgue measure of each subset $d_N(U_i \cap C(d_N))$ is zero. Thus the claim of Theorem 1 is clear. Fix $p \in C(d_N)$.

Since N is a C^∞ -smooth submanifold of M , there exists $\epsilon > 0$ small enough such that the set $N^\epsilon := \{x \in M : d_N(x) = \epsilon\}$ is a C^∞ -smooth hypersurface of M , and such that the distance function $d_N(\cdot)$ is of class C^∞ on the set $\{x \in M : d_N(x) \in (0, 2\epsilon)\}$. For such an $\epsilon > 0$, if $\gamma : [0, a] \rightarrow M$ is a N -segment joining N to $p = \gamma(a)$ with $a > \epsilon$, then the unit speed geodesic segment $\tilde{\gamma} : [\epsilon, a] \rightarrow M$ is a N^ϵ -segment joining N^ϵ to $p = \gamma(a)$. In particular, this means that the distance functions $d_N(\cdot)$ and d_{N^ϵ} have the same critical points and that

$$\forall p \in \{x \in M : d_N(x) > \epsilon\}, \quad d_N(x) = d_{N^\epsilon}(x) + \epsilon. \quad (20)$$

By construction, N^ϵ is a hypersurface with one normal vector at $q \in N^\epsilon$ given by $\nabla d_N(q)$ of norm 1. Hence, by completeness of M , for every $q \in N^\epsilon$ there exists a unique unit speed geodesic $\gamma_q(\cdot) : [0, \infty) \rightarrow M$ starting at q

with $\dot{\gamma}_q(0) = \nabla d_N(q)$. The distance from N^ϵ to the cut point on γ_q is given by

$$\rho(q) := \sup\{t : \gamma_q|_{[0,t]} \text{ is a } N^\epsilon\text{-segment}\}.$$

The function $\rho : N^\epsilon \rightarrow (0, \infty]$ is called the distance function from N^ϵ to the cut locus. In [7], Itoh and Tanaka proved the following (we refer the reader to [7] for the definition of the distance function to the cut locus $\rho(\cdot)$ whenever N is an arbitrary submanifold of M).

Theorem 6. *Let N be a C^∞ -smooth submanifold of a complete C^∞ -smooth Riemannian manifold M and $\pi : U\nu \rightarrow N$ the unit normal bundle of N . Then for each $v \in U\nu$ with $\rho(v) < \infty$, ρ is locally Lipschitz around v .*

By the discussion above $p \in C(d_N^\epsilon)$. Define the set $Q(p) \subset N$ by

$$Q(p) := \{q \in N^\epsilon : d(p, q) = d_{N^\epsilon}(p)\}.$$

The set $Q(p)$ is clearly compact and nonempty. Hence by Theorem 6, there exist some neighbourhood \mathcal{V} of $Q(p)$ in M such that $\rho : \mathcal{V} \rightarrow (0, \infty)$ is Lipschitz continuous on \mathcal{V} .

On the other hand, there exists a uniform normal neighbourhood \mathcal{W} of p and $\alpha > 0$ such that \mathcal{W} is contained in a geodesic ball of radius α around each of its points (see for instance [9, Lemma 5.12 p. 78] for such a result). In particular, this means that the function

$$\begin{aligned} d(\cdot, \cdot) : \mathcal{W} \times \mathcal{W} &\longrightarrow \mathbb{R} \\ (p_1, p_2) &\longmapsto d(p_1, p_2) \end{aligned}$$

is of class C^∞ on $\{(p_1, p_2) \in \mathcal{W} \times \mathcal{W} : p_1 \neq p_2\}$.

Set $d := d_{N^\epsilon}(p)$. For each element q of $Q(p)$, $\rho(q) = d$. Moreover by construction of N^ϵ for each element q of $Q(p)$, the unique unit speed geodesic starting at q and orthogonal to N^ϵ is minimizing on $[0, d]$. Hence there exists some neighbourhood $\mathcal{V}' \subset \mathcal{V}$ of $Q(p)$ such that $\overline{\mathcal{V}' \cap N^\epsilon}$ is a compact C^∞ -smooth hypersurface with boundary, such that the closure of the $(d - \alpha/2)$ -front $F(d - \alpha/2)$ of $N^\epsilon \cap \mathcal{V}'$ (that is, the closure of the set of $\gamma_q(d - \alpha/2)$ for any $q \in \mathcal{V}'$) is included in \mathcal{W} , and such that

$$\forall q \in \mathcal{V}', \quad \rho(q) \geq d - \frac{\alpha}{4}. \quad (21)$$

Consequently, by (21), the front $\overline{F(d - \alpha/2)}$ is a compact C^∞ -smooth hypersurface with boundary of M , which is included in \mathcal{W} . Moreover there exists some neighbourhood U of p which does not intersect $\overline{F(d - \alpha/2)}$ such that for every $p' \in U$,

$$d_N(p') = \epsilon + d - \frac{\alpha}{2} + d_{\overline{F(d - \alpha/2)}}(p').$$

We deduce that $d_N(\cdot)$ and $d_{\overline{F(d - \alpha/2)}}(\cdot)$ have locally the same critical points and that $d_{\overline{F(d - \alpha/2)}}(\cdot)$ is defined on U by

$$d_{\overline{F(d - \alpha/2)}}(p') = \min\{d(p', q) : q \in \overline{F(d - \alpha/2)}\}.$$

Since the function $d(\cdot, \cdot)$ is of class C^∞ outside the diagonal of \mathcal{W} , we are in the situation³ of Theorem 3. Moreover by [14, Proposition 4.8 p.108], we

³Here the front is a smooth compact submanifold with boundary of M . Theorem 3 holds in that case; basically no changes are required in the proof.

conclude that both definitions of critical points of $d_{F(d-\alpha/2)}$ coincide. Hence we conclude.

3.3. Proof of Theorem 2. In the Euclidean case, by (1) we have,

$$\partial d_N(x) = \text{co} \left\{ \frac{x - q}{\|x - q\|} : q \in N \right\}.$$

Hence we deduce that $x \notin N$ is a critical point of $d_N(\cdot)$ if and only if it is a critical point of $d_N(\cdot)^2$, that is if

$$0 \in \text{co}\{x - q : q \in N\}.$$

We conclude that

$$d_N(C(d_N)) = \{\sqrt{\rho} : \rho \in d_N^2(C(d_N^2))\}.$$

Since the function $\rho \mapsto \sqrt{\rho}$ is locally Lipschitz on $(0, \infty)$, it suffices to show that the set of all critical values of $d_N(\cdot)^2$ has Lebesgue measure zero.

As in the proof of Lemma 3, if we consider N as a submanifold of \mathbb{R}^m (by Whitney embedding's theorem), we can approximate N by a finite union of some algebraic sets $\mathcal{E}_{i,j}$ with $i \in I$ and $j \in \{1, \dots, l^d\}$. So if we call configuration, some $(2n+2)$ -tuple $(i_1, \dots, i_{n+1}, j_1, \dots, j_{n+1})$ with each i_s ($s = 1, \dots, n+1$) belonging to I , and with each j_s ($s = 1, \dots, n+1$) belonging to the set $\{1, \dots, n+1\}$, we enumerate $|I|^{n+1}l^{d(n+1)}$ such configurations. We have the following lemma.

Lemma 3. *There exists two positive constants M_1, M_2 such that if x belongs to $C(d_N(\cdot)^2) \cap \mathcal{T}$, then there exists some configuration \mathcal{S} such that x is a $(M_1 r^k, M_2 r^{k-1})$ -critical point of the mapping*

$$\begin{aligned} d_{\mathcal{S}}^2 : \mathcal{T} &\longrightarrow \mathbb{R} \\ x &\longmapsto \min\{\|x - q_s\|^2 : q_s \in \mathcal{E}_{i_s, j_s}, s \in \{1, \dots, n+1\}\}. \end{aligned}$$

Moreover we have,

$$|d_N(x)^2 - d_{\mathcal{S}}^2(x)| \leq M_1 r^k. \quad (22)$$

Proof. In this case the function $\phi(\cdot, \cdot)$ is already a polynomial, hence we do not have to approximate it by its Taylor polynomial. The calculations are easier than in Lemma 2 above. The proof is let to the reader. \square

Returning to the proof of Theorem 2. By the same arguments as in the proof of Lemma 3, we deduce that there exists two constants N and K such that for r small enough,

$$\begin{aligned} \text{meas}(d_N^2(C(d_N^2)) \cap \mathcal{T}) &\leq |I|^{n+1} l^{d(n+1)} 2K_P \sqrt{\frac{2M_1}{M_P}} \frac{1}{\sqrt{l^k}} \\ &\leq |I|^{n+1} 2K_P \sqrt{\frac{2M_1}{M_P}} l^{d(n+1) - \frac{k}{2}}. \end{aligned}$$

As l tends to infinity, since $k > 2(n+1)d$, we deduce that

$$\text{meas} f(C(f) \cap \mathcal{T}) = 0,$$

which conclude the proof of Theorem 2.

REFERENCES

- [1] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.
- [2] F.H. Clarke. Generalized gradients and applications. *Trans. Amer. Math. Soc.*, 205:247–262, 1975.
- [3] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, and P.R. Wolenski. *Nonsmooth Analysis and Control Theory*. Graduate Texts in Mathematics, vol. 178. Springer-Verlag, New York, 1998.
- [4] M. Coste. Ensembles semi-algébriques. In *Real algebraic geometry and quadratic forms (Rennes, 1981)*, volume 959 of *Lecture Notes in Math.*, pages 109–138. Springer, Berlin, 1982.
- [5] S. Ferry. When ϵ -boundaries are manifolds. *Fund. Math.*, 90(3):199–210, 1975/76.
- [6] J. H. G. Fu. Tubular neighborhoods in Euclidean spaces. *Duke Math. J.*, 52(4):1025–1046, 1985.
- [7] J. Itoh and M. Tanaka. The Lipschitz continuity of the distance function to the cut locus. *Trans. Amer. Math. Soc.*, 353(1):21–40, 2001.
- [8] J. Itoh and M. Tanaka. A Sard theorem for the distance function. *Math. Ann.*, 320(1):1–10, 2001.
- [9] J. M. Lee. *Riemannian manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. An introduction to curvature.
- [10] S. Lojasiewicz. Ensembles semi-analytiques. IHES Lecture notes. 1965.
- [11] A. P. Morse. The behavior of a function on its critical set. *Annals of Math.*, 40:62–70, 1939.
- [12] H. Rademacher. Über partielle und totale differenzierbarkeit von funktionen mehrerer variablen und über die transformation der doppelintegrale. *Math. Ann.*, 79:340–359, 1919.
- [13] R. T. Rockafellar. *Convex analysis*. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.
- [14] T. Sakai. *Riemannian geometry*, volume 149 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1996. Translated from the 1992 Japanese original by the author.
- [15] A. Sard. The measure of the critical values of differentiable maps. *Bull. Amer. Math. Soc.*, 48:883–890, 1942.
- [16] S. Sternberg. *Lectures on differential geometry*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
- [17] H. Whitney. A function not constant on a connected set of critical points. *Duke Math. J.*, 1:514–517, 1935.
- [18] Y. Yomdin. The geometry of critical and near-critical values of differentiable mappings. *Math. Ann.*, 264(4):495–515, 1983.
- [19] Y. Yomdin. Metric properties of semialgebraic sets and mappings and their applications in smooth analysis. In *Géométrie algébrique et applications, III (La Rabida, 1984)*, volume 24 of *Travaux en Cours*, pages 165–183. Hermann, Paris, 1987.