The Quantitative Morse Theorem

Ta Lê Loi

University of Dalat, Dalat, Vietnam
loitl@dlu.edu.vn

Phan Phien

Nhatrang College of Education, Nhatrang, Vietnam
phieens@yahoo.com

Abstract

In this paper, we give a proof of the quantitative Morse theorem stated by Y. Yomdin in [12]. The proof is based on the quantitative Sard theorem, the quantitative inverse function theorem and the quantitative Morse lemma.

Mathematics Subject Classification: Primary 58K05; Secondary 58E05, 97N40

Keywords: Morse theorem, Quantitative assessment, Critical and near-Critical point

1 Introduction

One of the first basic results of classical singularity theory are that Sard Theorem [11] and [7], and Morse theorem [6]. These theorems research critical points and critical values of smooth mappings on open subsets of $\mathbb{R}^n$. The quantitative assessments and applications of the theorems were also considered. Y. Yomdin in [13] introduced the concept of near-critical points and near-critical values of a map, and there have been many results on quantitative assessments for the set of these points and values. One of them is the quantitative Sard theorem for mappings of class $C^k$ (see [12], [13], [14], [15] and [10]). The results give some explicit bounds in term of $\varepsilon$-entropy of the set of near-critical values.

This research is supported by Vietnam’s National Foundation for Science and Technology Development (NAFOSTED).
For Morse theorem, in [12] Y. Yomdin also stated a quantitative form for $C^k$-functions. But in the article, he just gave a few suggestions without details for the proof of the theorem. Up to now, he probably hasn’t published the proof.

In this paper, we give a detailed proof of the quantitative Morse theorem. The proof is based on the quantitative Sard theorem, the quantitative inverse function theorem and the quantitative Morse lemma.

2 Preliminaries

We give here some definitions, notations and results that will be used later.

Let $M_{m \times n}$ denote the vector space of real $m \times n$ matrices,

$$
\|x\| = (|x_1|^2 + \cdots + |x_n|^2)^{\frac{1}{2}}, \text{ where } x = (x_1, \ldots, x_n) \in \mathbb{R}^n,
$$

$B^n$ denotes the unit ball in $\mathbb{R}^n$, $B^n_r$ denotes the ball of radius $r$, centered at $0 \in \mathbb{R}^n$, and $B^n_r(x_0)$ denotes the ball of radius $r$, centered at $x_0 \in \mathbb{R}^n$,

$$
\|A\| = \max_{\|x\|=1} \|Ax\|, \text{ where } A \in M_{m \times n},
$$

$$
\|A\|_{\text{max}} = \max_{i,j} |a_{ij}|, \text{ where } A = (a_{ij})_{m \times n} \in M_{m \times n},
$$

$B_{n \times n}$ denotes the unit ball in $M_{n \times n}$, $\text{Sym}(n)$ denotes the space of real symmetric $n \times n$-matrices.

**Definition 2.1.** Let $f : M \to \mathbb{R}^m$ be a differentiable mapping class $C^k$, $M \subset \mathbb{R}^n$. Then $C^k$-norm of $f$ is defined by

$$
\|f\|_{C^k} = \sum_{j=1}^{k} \sup_{x \in M} \|D^j f(x)\|.
$$

**Definition 2.2.** A mapping $f : \mathbb{R}^m \to \mathbb{R}^n$ is called Lipschitz in a neighborhood of a point $x_0$ in $\mathbb{R}^n$ if there exists a constant $K > 0$ such that for all $x$ and $y$ near $x_0$, we have

$$
\|f(x) - f(y)\| \leq K \|x - y\|.
$$

Then we call $f$ the $K$-Lipschitz.

The usual $m \times n$ Jacobian matrix of partial derivatives of $f$ at $x$, when it exists, is denoted by $Jf(x)$. By Rademacher’s theorem (see [3, Theorem 3.1.6]), we have the following definition:
Definition 2.3 (F. H. Clarke - [C1], [C2]). The generalized Jacobian of $f$ at $x_0$, denoted by $\partial f(x_0)$, is the convex hull of all matrices $M$ of the form

$$M = \lim_{i \to \infty} J_f(x_i),$$

where $f$ is differentiable at $x_i$ and $x_i$ converges to $x_0$ for each $i$. $\partial f(x_0)$ is said to be of maximal rank if every $M$ in $\partial f(x_0)$ is of maximal rank.

Theorem 2.4 (Quantitative inverse function theorem, c.f. [1] and [9]). Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a $K$-Lipschitz mapping in a neighborhood of a point $x_0$ in $\mathbb{R}^n$. Suppose that $\partial f(x_0)$ is of maximal rank, set

$$\delta = \frac{1}{2} \inf_{M \in \partial f(x_0)} \frac{1}{\|M^{-1}\|},$$

$r$ be chosen so that $f$ satisfies $K$-Lipschitz condition and $\partial f(x) \subset \partial f(x_0) + \delta \mathbb{B}_{n \times n}$, when $x \in \mathbb{B}_{r(x_0)}^n$. Then $f$ is invertible in $\mathbb{B}_{\frac{\delta}{2}}^n(f(x_0))$ and there exists the inverse mapping

$$g : \mathbb{B}_{\frac{\delta}{2}}^n(f(x_0)) \to \mathbb{R}^n$$

being $\frac{1}{\delta}$-Lipschitz.

Definition 2.5 (Singular values of linear mapping, c.f. [4]). Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. Then there exist $\sigma_1(L) \geq \ldots \geq \sigma_r(L) > 0$, where $r = \text{rank } L$, so that $L(\mathbb{B}^n)$ is an $r$-dimensional ellipsoid of semi-axes $\sigma_1(L) \geq \ldots \geq \sigma_r(L)$. Set $\sigma_0(L) = 1$ and $\sigma_{r+1}(L) = \ldots = \sigma_m(L) = 0$, when $r < m$. We call $\sigma_0(L), \ldots, \sigma_m(L)$ the singular values of $L$.

Remark 2.6. Let $L$ be a linear mapping or a matrix. Then

(i) $\sigma_{\text{max}}(L) = \|L\| = \sigma_1(L), \quad \sigma_{\text{min}}(L) = \min_{\|x\|=1} \|Lx\|$.

(ii) When $L \in L(\mathbb{R}^n, \mathbb{R}^n)$, and $\lambda$ is an eigenvalue of $L$, we have

$$\sigma_{\text{min}}(L) \leq |\lambda| \leq \sigma_{\text{max}}(L).$$

Definition 2.7. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a $k$ times differentiable mapping, $k \geq 1$. For $\Lambda = (\lambda_1, \ldots, \lambda_m)$, $\lambda_1 \geq \ldots \geq \lambda_m \geq 0$, we call

$$\Sigma(f, \Lambda) = \{x \in \mathbb{R}^n : \sigma_i(Df(x)) \leq \lambda_i, i = 1, \ldots, m\}$$

the set of $\Lambda$-critical points of $f$, and

$$\Delta(f, \Lambda) = f(\Sigma(f, \Lambda))$$
the set of \textit{\Lambda}-critical values of \( f \).

Set \( \Sigma(f, \Lambda, A) = \Sigma(f, \Lambda) \cap A \), \( \Delta(f, \Lambda, A) = f(\Sigma(f, \Lambda, A)) \), \( A \subset \mathbb{R}^n \). When \( \gamma = (\gamma, \ldots, \gamma) \in \mathbb{R}_+^m \), a point \( y \in \mathbb{R}^m \) is called \( \gamma \)-regular value of \( f \) if \( y \notin \Delta(f, \gamma, A) \), i.e. \( f^{-1}(y) = \emptyset \) or if \( x \in f^{-1}(y) \) then there exists a number \( i \in \{1, \ldots, m\} \) so that \( \sigma_i(Df(x)) \geq \gamma \).

\textbf{Remark 2.8.} If \( \Lambda = (0, \ldots, 0) \) then \( \Sigma(f, 0) \) is the set of critical points and \( \Delta(f, 0) \) is the set of critical values of \( f \).

\textbf{Definition 2.9.} Let \( X \) be a metric space, \( A \subset X \) a relatively compact subset. For any \( \varepsilon > 0 \), denoted by \( M(\varepsilon, A) \) the minimal number of closed balls of radius \( \varepsilon \) in \( X \), covering \( A \).

\textbf{Theorem 2.10} (Quantitative Sard theorem, c.f. \cite[Theorem 9.6]{15}). Let \( f : B_r^n \rightarrow \mathbb{R}^m \) be a mapping of class \( C^k \), \( q = \min(n, m) \), \( \Lambda = (\lambda_1, \ldots, \lambda_q) \), \( \lambda_i > 0, i = 1, \ldots, q \), \( B_\delta \) is a ball of radius \( \delta \) in \( \mathbb{R}^m \). When \( 0 < \varepsilon \leq \delta \)

\[
M(\varepsilon, \Delta(f, \Lambda, B_r^n \cap B_\delta^m)) \leq c \left( \frac{R_k(f)}{\varepsilon} \right)^{\frac{q}{k}} \sum_{i=0}^{q} \min \left( \lambda_0 \ldots \lambda_i \left( \frac{r}{\varepsilon} \right)^i \left( \frac{\varepsilon}{R_k(f)} \right)^i \left( \frac{\delta}{\varepsilon} \right)^i \right),
\]

where \( c = c(n, m, k) \), \( R_k(f) = \frac{K}{(k-1)!} r^{k-1} \), and \( K \) is a Lipschitz constant of \( D^{k-1}f \) in \( B_r^n \).

\textbf{Lemma 2.11} (Quantitative Morse lemma). Let \( A \in \text{Sym}(n) \). Suppose that \( Q_0 \in \text{Gl}(n) \) such that \( tQ_0AQ_0 = D_0 = \text{diag}(1, \ldots, 1, -1, \ldots, -1) \). Set

\[
U(A) = \{ B \in \text{Sym}(n) : \|B - A\| \leq \frac{1}{2n\|Q_0\|^2} \}.
\]

Then there exists a mapping \( P : U(A) \rightarrow \text{Gl}(n) \subset C^\omega \) satisfying

\[
P(A) = Q_0, \qquad \text{and if } P(B) = Q \text{ then } tQBQ = D_0.
\]

\textbf{Proof.} For \( B \in U(A) \), we have

\[
\|tQ_0BQ_0 - tQ_0AQ_0\|_{\text{max}} \leq \|tQ_0(B - A)Q_0\| \leq \|Q_0\|^2 \|B - A\| \leq \frac{1}{2n}.
\]

If \( tQ_0BQ_0 = (b_{ij})_{1 \leq i, j \leq n} \) then \( |b_{ii}| \geq \sum_{j \neq i} |b_{ij}|. \) So \( \det(b_{ij})_{1 \leq i, j \leq k} \neq 0 \), for \( k = 1, \ldots, n \). Therefore, the normalization (see \cite[Lemma p.145]{5}) reductioning \( tQ_0BQ_0 \) to the normal form \( D_0 \) defines the mapping \( P : U(A) \rightarrow \text{Gl}(n) \subset C^\omega \) satisfying the demands of the lemma. \( \square \)
Remark 2.12. The reduction a non-degenerate real symmetric matrix $A$ to the normal form $D_0$ can be realized by a matrix $Q_0$ of the form $Q_0 = SU$, where $U$ is a orthogonal matrix, and $S$ is a diagonal matrix. So

$$\|Q_0\|^2 = \frac{1}{\sigma_{\min}(A)}.$$

3 The quantitative Morse theorem

Theorem 3.1 (c.f. [12, Theorem 4.1, Theorem 6.1]). Fix $k \geq 3$. Let $f_0 : \mathbb{B}^n \rightarrow \mathbb{R}$ be a $C^k$-function in a open set contain $\mathbb{B}^n$ with all derivatives up to order $k$ uniformly bounded by $K$. Then for any given $\varepsilon > 0$, we can find $h$ with $\|h\|_{C^k} \leq \varepsilon$ and the positive functions $\psi_1$, $\psi_2$, $\psi_3$, $d$, $M$, $N$, $\eta$ depending on $K$ and $\varepsilon$, such that $f = f_0 + h$ satisfies the following conditions:

(i) At each critical point $x_i$ of $f$, the smallest absolute value of the eigenvalues of the Hessian $Hf(x_i)$ is at least $\psi_1(K, \varepsilon)$.

(ii) For any two different critical points $x_i$ and $x_j$ of $f$, $\|x_i - x_j\| \geq d(K, \varepsilon)$. Consequently, the number of the critical points does not exceed $N(K, \varepsilon)$.

(iii) For any two different critical points $x_i$ and $x_j$ of $f$, $|f(x_i) - f(x_j)| \geq \psi_2(K, \varepsilon)$.

(iv) For $\delta = \psi_3(K, \varepsilon)$ and for each critical point $x_i$ of $f$, there exists a coordinate transformation $\varphi : \mathbb{B}^n_\delta(x_i) \rightarrow \mathbb{R}^n \in C^r$ such that

$$f \circ \varphi^{-1}(y_1, \ldots, y_n) = y_1^2 + \cdots + y_i^2 - y_{i+1}^2 - \cdots - y_n^2 + \text{const},$$

and $\|\varphi\|_{C^{k-1}} \leq M(K, \varepsilon)$.

(v) If $\|\nabla f(x)\| \leq \eta(K, \varepsilon)$, then $x \in \mathbb{B}^n_\delta(x_i)$, with $x_i$ is a critical point of $f$.

The proof of (i) is based on the suggestion of Y. Yomdin (see [14]). The proofs of (ii), (iii), (iv) and (v) are based on the quantitative inverse theorem and the quantitative Morse lemma in section 2.

Proof. 

(i) Let $\varepsilon > 0$. Applying Theorem 2.10,

$$M(r, \Delta(Df_0, \gamma, \mathbb{B}^n) \cap \mathbb{B}^n_\varepsilon) \leq cR_k(f_0)^\frac{1}{2} \sum_{i=0}^{n} \left(\frac{\gamma}{R_k(f_0)^i r^{\frac{1}{2i}}}\right)^i.$$
When $r < 1$ and $\gamma < rR_k(f_0)^{\frac{1}{k}}$,

$$M(r, \Delta(Df_0, \gamma, \overline{B}^n) \cap B^n_\varepsilon) \leq cR_k(f_0)^{\frac{n}{k}} \frac{1}{r^{\frac{n}{k}}} \sum_{i=0}^{n} \left( \frac{\gamma}{rR_k(f_0)^{\frac{n}{k}}} \right)^i \leq cR_k(f_0)^{\frac{n}{k}} \frac{1}{r^{\frac{n}{k}}} \frac{1}{1 - \frac{\gamma}{rR_k(f_0)^{\frac{n}{k}}}}.$$ 

So the Lebesgue measure of $\Delta(Df_0, \gamma, \overline{B}^n) \cap B^n_\varepsilon$,

$$m(\Delta(Df_0, \gamma, \overline{B}^n) \cap B^n_\varepsilon) \leq r^n m(\overline{B}^n) M(r, \Delta(Df_0, \gamma, \overline{B}^n) \cap B^n_\varepsilon) \leq r^n m(\overline{B}^n) cR_k(f_0)^{\frac{n}{k}} \frac{1}{r^{\frac{n}{k}}} \frac{1}{1 - \frac{\gamma}{rR_k(f_0)^{\frac{n}{k}}}}.$$ 

Let

$$r(\varepsilon) = \frac{1}{2} \min(\varepsilon, \left(\frac{\varepsilon}{c^n R_k(f_0)^{\frac{n}{k}}}\right)^{\frac{1}{k-1}}),$$

and

$$\gamma(K, 2\varepsilon) = R_k(f_0)^{\frac{1}{k}} r(\varepsilon)(1 - \frac{r(\varepsilon)^n cR_k(f_0)^{\frac{n}{k}}}{\varepsilon^n r(\varepsilon)^{\frac{n}{k}}}) > 0,$$

with $R_k(f_0) = \frac{K}{(k-1)!}$. We get

$$m(\Delta(Df_0, \gamma, \overline{B}^n) \cap B^n_\varepsilon) < \varepsilon^n m(\overline{B}^n) = m(B^n_\varepsilon).$$

So we can choose a $\gamma(K, 2\varepsilon)$-regular value $v$ of $Df_0$, with $\|v\| < \varepsilon$.

Now, let $h : \overline{B}^n \to \mathbb{R}$ be a linear mapping with $Dh = -v$ and $f = f_0 + h$. Then $\|h\|_c \leq \varepsilon, Df = Df_0 - v$, and $Hf = Hf_0 = D(Df_0)$. So $0$ is a $\gamma(K, 2\varepsilon)$-regular value of $Df$, and at each critical point $x_i$ of $f$, we have

$$\|Hf(x_i)\| \geq \gamma(K, 2\varepsilon). \quad (3.1)$$

By Remark 2.6, the smallest absolute value of the eigenvalues of the Hessian $Hf(x_i)$ is at least $\psi_1(K, 2\varepsilon) = \gamma(K, 2\varepsilon)$.

(ii) Consider $Df : \overline{B}^n \to \mathbb{R}^n$. Suppose that $x_i$ is a critical point of $f$. Then applying (3.1) we obtain

$$\delta' = \frac{1}{2} \frac{1}{\|Hf(x_i)\|^{-1}} \geq \frac{1}{2} \gamma(K, 2\varepsilon).$$

Choose $\delta' = \frac{1}{2} \gamma(K, 2\varepsilon)$, we have

$$\|D(Df)(x) - D(Df)(x_i)\| = \|D(Df_0)(x) - D(Df_0)(x_i)\| \leq K\|x - x_i\|,$$

hence, if $\|x - x_i\| \leq \frac{\delta'}{K}$, we get

$$D(Df)(x) \in D(Df)(x_i) + \delta' B_{n \times n}.$$
Therefore, if \( r = \frac{\delta'}{K} \), then every \( x \in B^n_r(x_i) \) we obtain

\[
D(Df)(x) \in D(Df)(x_i) + \delta'B_{n \times n}.
\]

Thus, applying Theorem 2.4, \( Df \) is invertible in 

\[
B^n_{\frac{\delta'}{2K}}(x_i) = B^n_{\frac{\delta'(K,\varepsilon)}{8K^2}}(x_i).
\]

Hence, \( Df^{-1}(0) \) is unique in \( B^n_{\frac{\delta'(K,\varepsilon)}{8K^2}}(x_i) \), i.e. \( x_i \) is the unique critical point of \( f \) in the ball \( B^n_{\frac{\delta'(K,\varepsilon)}{8K^2}}(x_i) \).

So if \( x_i, x_j \) are different critical points of \( f \), we have

\[
d(x_i, x_j) \geq d(K, 2\varepsilon) = \frac{1}{4} \frac{\gamma^2(K, 2\varepsilon)}{K^2} > 0.
\]

Therefore, the number of critical points \( x_i \) does not exceed

\[
N(K, 2\varepsilon) = M \left( \frac{1}{4} \frac{\gamma^2(K, 2\varepsilon)}{K^2}, B^n \right).
\]

(iii) Suppose that the number of critical points of \( f \) being \( N, N \leq N(K, 2\varepsilon) \), and critical values of \( f \) ordered as follows:

\[
f(x_1) \leq f(x_2) \leq \ldots \leq f(x_N).
\]

For each critical point \( x_i \) of \( f \), set

\[
U_i = B^n_{\frac{\delta'(K,\varepsilon)}{2K}}(x_i) \cap \overline{B^n}, \quad B_i = B^n_{\frac{\delta'(K,\varepsilon)}{4K^2}}(x_i) \cap \overline{B^n}.
\]

We call \( \lambda_i : \overline{B^n} \rightarrow [0, 1] \) the mapping of class \( C^\infty \), where

\[
\lambda_i(x) = \begin{cases} 0, & x \notin U_i \\ 1, & x \in B_i \end{cases}
\]

with all derivatives uniformly bounded by \( C_1 \). Set \( \tilde{f} = f + \lambda \), with

\[
\lambda : \overline{B^n} \rightarrow \mathbb{R}, \quad \lambda(x) = \sum_{i=1}^{N} c_i \lambda_i(x), \text{ where } c_i = i \cdot \frac{\varepsilon}{2C_1kN^2} > 0.
\]

From (ii) we obtain every \( U_i \) disjoint, and we have \( \| \lambda \|_{C^k} \leq \tilde{\varepsilon} \).

Thus \( \tilde{f} \) will be a Morse function having the same critical points as \( f \) and these will have the same indices. Moreover, \( \tilde{f}(x_i) = f(x_i) + c_i \). Hence, with \( x_i, x_j \) are critical points, \( i \neq j \), we obtain

\[
|\tilde{f}(x_i) - \tilde{f}(x_j)| = |f(x_i) + c_i - f(x_j) - c_j| \geq \frac{\varepsilon}{2kC_1N^2} > 0. \quad (3.2)
\]
Therefore, replacing the linear mapping \( h \) in \((i)\) by 

\[
h = h_1 + \lambda,
\]

with \( h_1 : \mathbf{B}^n \to \mathbb{R} \) being a linear mapping such that \( Dh_1 = -v \), and \( v \) is a \( \gamma(K, \varepsilon) \)-regular value of \( Df_0 \), at a distance of most \( \frac{\varepsilon}{2} \) from 0, we get

\[
\|h\|_{C^k} = \|h_1 + \lambda\|_{C^k} \leq \varepsilon,
\]

and \( f = f_0 + h = f_0 + h_1 + \lambda \) to satisfy \((i)\) and \((ii)\), with

\[
\psi_1(K, \varepsilon) = \gamma(K, \varepsilon); \quad d(K, \varepsilon) = \frac{1}{4} \frac{\gamma^2(K, \varepsilon)}{K^2}, \quad N(K, \varepsilon) = M \left( \frac{1}{4} \frac{\gamma^2(K, \varepsilon)}{K^2}, \mathbf{B}^n \right).
\]

Moreover, by \((3.2)\), for any \( i \neq j \), we have

\[
|f(x_i) - f(x_j)| \geq \psi_2(K, \varepsilon) = \frac{\varepsilon}{2kC_1N^2(K, \varepsilon)} > 0.
\]

\((iv)\) According to \((ii)\), we only need to prove \((iv)\) for each critical point \( x_i \).

Moreover, we may assume \( x_i = 0 \), \( f(x_i) = 0 \).

Let \( Q_0 \in \text{Gl}(n) \) be a linear transformation satisfying the condition of Remark 2.12 such that

\[
^tQ_0 H f(0) Q_0 = D_0.
\]

The coordinate transformation \( \varphi \) is constructed as follows.

First, let \( \mathcal{B} : \mathbf{B}^n \to \text{Sym}(n) \in C^{k-1} \) in a open set contain \( \mathbf{B}^n \), be defined by

\[
\mathcal{B}(x) = B_x = (b_{ij}(x))_{1 \leq i, j \leq n},
\]

where

\[
b_{ij}(x) = \int_0^1 \int_0^1 \frac{\partial^2 f}{\partial x_j \partial x_i}(stx)dsdt, 1 \leq i, j \leq n.
\]

Then

\[
f(x) = \sum_{i,j=1}^n b_{ij}(x)x_i x_j \text{ and } \mathcal{B}(0) = A = H f(0).
\]

Applying Lemma 2.11, we get

\[
\mathcal{P} : U(A) \to \text{Gl}(n),
\]

being of class \( C^\omega \) such that \( \mathcal{P}(A) = Q_0 \), and if \( \mathcal{P}(B) = Q \) then \(^tQBQ = D_0\).

According to the Mean Value Theorem and Remark 2.12, the condition to apply Lemma 2.11 is

\[
\|H f(x) - A\| \leq (K + \varepsilon)\|x\| \leq \frac{1}{2n\|Q_0\|^2} = \frac{1}{2n}\sigma_{\text{min}}(A),
\]
or
\[ \|x\| \leq \frac{1}{(K + \varepsilon)2n\|Q_0\|^2} = \frac{1}{(K + \varepsilon)2n}\sigma_{\min}(A). \]
Set \( \delta = \psi_3(K, \varepsilon) = \frac{1}{(K + \varepsilon)2n}\gamma(K, \varepsilon) \) and
\[ \varphi : U_\delta(0) \to \mathbb{R}^n, \quad y = \varphi(x) = Q_x^{-1}x, \quad \text{with } Q_x = \mathcal{P}(B_x). \]
We have
\[ f(x) = t' x B_x x = t' y(Q_x B_x Q_x y) = t' y D_0 y = y_1^2 + \cdots + y_i^2 - y_{i+1}^2 - \cdots - y_n^2. \]
To prove \( \|\varphi\|_{C^{k-1}} \leq M(K, \varepsilon) \), present \( \varphi \) as the following composition
\[ \varphi : x \in U_\delta(0) \xrightarrow{B} B_x \xrightarrow{\mathcal{P}} Q_x \xrightarrow{\text{Inv}} Q_x^{-1} \xrightarrow{L} \varphi(x) = Q_x^{-1}x. \]
By the construction \( B \in C^{k-1} \), and by the assumption, the partial derivatives of \( B \)
\[ \|\partial^\alpha B(x)\| \leq K, \text{ for all } \alpha \in \mathbb{N}^n, |\alpha| \leq k - 1. \]
Since \( U(A) \) is compact, there exists \( M_1(K, \varepsilon) > 0 \) such that
\[ \|\partial^\alpha \mathcal{P}(B)\| \leq M_1(K, \varepsilon), \text{ for all } B \in U(A), |\alpha| \leq k - 1. \]
Similarly, since \( \mathcal{P}(U(A)) \) is compact, there exists \( C_2(K, \varepsilon) > 0 \) such that
\[ \|\partial^\alpha \text{Inv}(Q)\| \leq C_2(K, \varepsilon), |\alpha| \leq k - 1, \text{ for all } Q \in \mathcal{P}(U(A)). \]
Let
\[ \overline{L} : U_\delta \times \text{Inv}(\mathcal{P}(U(A))) \to \mathbb{R}^n, \overline{L}(x, Q') = Q' x. \]
Then \( \overline{L} \) is a bilinear form. Hence there exists \( C_3(K, \varepsilon) > 0 \) such that
\[ \|\partial L\| \leq C_3(K, \varepsilon), \|\partial^\alpha L\| = 0, \text{ for } |\alpha| \geq 2, \text{ and } (x, Q') \in U_\delta \times \text{Inv}(\mathcal{P}(U(A))). \]
Since \( \partial^\alpha \varphi \) can be represented as a sum of products of \( \partial^\alpha B, \partial^\alpha \mathcal{P}, \partial^\alpha \text{Inv} \) and
\( \partial^\alpha L \), with \( |\alpha_j| \leq |\alpha|, j = 1, \ldots, 4 \), there exists \( M(K, \varepsilon) > 0 \) depending on \( K \),
\( M_1(K, \varepsilon), C_2(K, \varepsilon) \) and \( C_3(K, \varepsilon) \) such that \( \|\varphi\|_{C^{k-1}} \leq M(K, \varepsilon). \)
(v) Consider \( Df : \mathbb{B}^n \to \mathbb{R}^n. \) Then for \( x_i \) is critical point of \( f \), we have
\[ Df(x_i) = 0, \|Df(x)\| = \|Df(x) - Df(x_i)\| \text{ for all } x \in \mathbb{B}^n, \]
moreover
\[ \sigma = \frac{1}{2} \frac{1}{\|Hf(x_i)\|} \geq \frac{1}{2} \gamma(K, \varepsilon). \]
We have
\[ \|D(Df)(x) - D(Df)(x_i)\| \leq (K + \varepsilon)\|x - x_i\|. \]
hence with $\|x - x_i\| \leq \frac{\sigma}{K + \varepsilon}$, and $\sigma = \frac{1}{2} \gamma(K, \varepsilon)$, we obtain

$$D(Df)(x) \in D(Df)(x_i) + \sigma B_{n \times n}.$$  

Therefore, if $r = \min(\frac{\sigma}{K + \varepsilon}, \frac{1}{\sigma_n} \gamma(K, \varepsilon))$, then

$$D(Df)(x) \in D(Df)(x_i) + \sigma B_{n \times n}, \text{ for all } x \in B^n_r(x_i).$$  

Hence, applying Theorem 2.4 there exist neighborhoods $U$ and $V$ of $x_i$ and $Df(x_i)$, respectively, such that $Df$ is invertible, with

$$U = B^n_{\frac{r \sigma}{2(K + \varepsilon)}}(x_i), \quad V = B^n_{\frac{r \sigma}{2}}(0).$$

So, with $\eta(K, \varepsilon) = \frac{r \sigma}{2} = \frac{1}{4} r \gamma(K, \varepsilon)$, as $\|\text{grad} f(x)\| \leq \eta(K, \varepsilon)$ we have

$$x \in B^n_{\frac{r \sigma}{2(K + \varepsilon)}}(x_i) \subset B^n_{\psi \gamma(K, \varepsilon)} (x_i).$$

□

References

[1] F. H. Clarke, On the inverse function theorem, Pacific Journal of Mathematics, Vol 64, No 1 (1976), 97-102.

[2] F. H. Clarke, Generalized gradients and applications, Trans. Amer. Math. Soc., Vol 205 (1975), 247-262.

[3] H. Federer, Geometric measures theory, Springer-Verlag, 1969.

[4] G. H. Golub and C. F. van Loan, Matrix computation, Johns Hopkins Univ. Press 1983.

[5] M. W. Hirsch, Differential Topology, Springer-Verlag, New York - Heidelberg - Berlin, 1976.

[6] M. Morse, The critical points of a function of $n$ variables, Trans. Amer. Math. Soc., 33, (1931), 71-91.

[7] A. Morse, The behavior of a function on its critical set, Ann. Math, 40 (1939), 62-70.

[8] L. Niederman, Prevalence of exponential stability among nearly integrable Hamiltonian systems, Ergodic Theory and Dynamical Systems, (2007), 25p.
[9] P. Phien, Some quantitative results on Lipschitz inverse and implicit functions theorems, East-West J. Math. Vol. 13, No 1 (2011), 7-22.

[10] A. Rohde, On the $\varepsilon$-Entropy of Nearly Critical Values, Journal of Approximation Theory, 76 (1994), 166-194.

[11] A. Sard, The measure of the critical values of differentiable maps, Bull. Amer. Math. Soc. 48, (1942), 883-890.

[12] Y. Yomdin, Some quantitative results in singularity theory, Anales Polonici Mathematici, 37 (2005), 277-299.

[13] Y. Yomdin, The Geometry of Critical and Near-Critical Values of Differentiable Mappings, Math. Ann. 264, (1983), 495-515.

[14] Y. Yomdin, Metric properties of semialgebraic sets and mappings and their applications in smooth analysis, (Proceedings of the Second International Conference on Algebraic Geometry, La Rabida, Spain, 1984, J.M. Aroca, T. Sahchez-Geralda, J.L. Vicente, eds.), Travaux en Cours, Hermann, Paris (1987), 165-183.

[15] Y. Yomdin and G. Comte, Tame geometry with application in smooth analysis, LNM vol. 1834, 2004.

Received: September, 2011