A SIMPLE PROOF OF THE MORSE-SARD THEOREM IN SOBOLEV SPACES

ALESSIO FIGALLI

(Communicated by David Preiss)

Abstract. In this paper we give a new simple proof of a result of Luigi De Pascale, which states that the Morse-Sard Theorem holds under the hypothesis of Sobolev regularity. Moreover, as our proof is independent of the Morse-Sard Theorem with $C^k$ regularity, our result implies the classical Morse-Sard Theorem.

The Morse-Sard Theorem is concerned with the size of the image of the critical values of a differentiable function. To recall it and to state our result, we need some definitions.

Definition 1. Let $\Omega \subset \mathbb{R}^n$ be open and let $f : \Omega \to \mathbb{R}^m$ be a $C^1$ function. A point $x \in \Omega$ is said to be a critical point if $Df(x)$ is not of maximum rank. A point $y \in f(\Omega)$ is said to be a critical value if $y = f(x)$ for a critical point $x$. The set of all the critical points is called the critical set.

Let us denote by $\mathcal{L}^m$ the $m$-dimensional Lebesgue measure. We can now recall the classical Morse-Sard Theorem (for a proof, see [1, Paragraph 15]):

Theorem 2 (Morse-Sard). Let $\Omega \subset \mathbb{R}^n$ be open and let $f : \Omega \to \mathbb{R}^m$ be a $C^{n-m+1}$ function, with $n \geq m$ ($C^1$ if $m > n$). Then the set of critical values of $f$ has $\mathcal{L}^m$-measure zero.

After that theorem, many generalizations have been proved and, at the same time, many counterexamples have been found in the case of not sufficient regularity. In particular, in [2] the same conclusion of the Morse-Sard Theorem has been proved under the only assumption of a $C^{n-m,1}$ regularity, while in [3] only a $W^{n-m+1,p}$ regularity, with $p > n$, is assumed (see [3] for more historical notes). Here we give a simple proof of the result in [3]. We remark that, as our proof is independent of Theorem 2, our result implies the classical Morse-Sard Theorem.

In the proof of our theorem, we will need a refined version of the classical Morrey inequality (for a proof, see [4, paragraph 4.5.3]):

Lemma 3. Let $\Omega \subset \mathbb{R}^n$ be an open subset and let $B(x,r)$ be a ball contained in $\Omega$. Then for any $y \in B(x,r)$ we have

\begin{equation}
|u(x) - u(y)| \leq C r^{1-\frac{n}{p}} \left( \int_{B(x,r)} |Du(z)|^p \, dz \right)^{\frac{1}{p}} \quad \forall u \in W^{1,p}.
\end{equation}
We remark that, in particular, this inequality gives the embedding $W^{1,p} \hookrightarrow C^{0,\alpha(p)}$, with $\alpha(p) = 1 - \frac{n}{p}$, and, in more generality, $W^{1,p} \hookrightarrow C^{1,1-\alpha(p)}$.

We will also need the Kneser-Glaeser Rough Composition Theorem. In order to state it, we recall that, given a positive integer $s$, a map $f$ is said to be $s$-flat on $A$ if $D^j f(x) = 0$ for $j = 1, \ldots, s$ for any $x \in A$.

**Theorem 4.** Let $W \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ be open sets; $A^* \subset W$ and $A \subset V$, with $A$ closed relative to $V$; $f : V \to \mathbb{R}^p$ of class $C^r$ on $V$ and $s$-flat on $A$; $g : W \to V$ of class $C^{r-s}$ with $g(A^*) \subset A$. Then there is a map $H : W \to \mathbb{R}^p$ of class $C^r$ satisfying:

(i) $H(x) = f(g(x))$ for $x \in A^*$;

(ii) $H$ is $s$-flat on $A^*$.

The proof of this theorem relies on Whitney’s Extension Theorem (see for example [1, Theorem 13.2]). Indeed, differentiating the identity $H = f \circ g$, one prescribes the derivative of $H$ on $A^*$, and then one only needs to check that the hypotheses needed to apply Whitney’s Theorem are satisfied (see [1, Theorem 14.1] for a detailed proof).

**Theorem 5.** Let $\Omega \subset \mathbb{R}^n$ be open and let $f : \Omega \to \mathbb{R}^m$ be a $W^{n-m+1,p}$ function, with $p > n \geq m$. Then the set of critical values of $f$ has $\mathcal{L}^m$-measure zero.

**Remark.** As $W^{n-m+1,p} \hookrightarrow C^{n-m,\alpha(p)}$, we will always refer to the $C^{n-m,\alpha(p)}$ representative. Moreover we observe that with the only assumption of $C^{n-m,\alpha}$ regularity with $\alpha < 1$ the result is false. The key point is in fact the existence of another weak derivative summable enough, as we will see in the proof.

**Proof.** First we observe that, as it suffices obviously to prove the theorem for $f$ restricted to each compact set of $\Omega$, we can assume that $\Omega$ is bounded and that $f \in W^{n-m+1,p}(\Omega, \mathbb{R}^m)$. Thanks to this remark, in the sequel we will always skip the subscript $\text{loc}$.

To simplify the notation, we define $k := n - m + 1$. We remark that, in the case $n = m$, the result is just a corollary of the area formula for Sobolev functions [1] (for a proof and for more references on the subject, see [5]), so we can assume $n > m$, that is, $k \geq 2$.

Let $C_f$ be the critical set of $f$ and let us define the sets

$$A_s := \{x \in \Omega \mid D^i f(x) = 0 \text{ for } 1 \leq i \leq s\}, \quad 1 \leq s \leq n - m,$$

and

$$K := \{x \in \Omega \mid 1 \leq \text{rank } Df(x) \leq m - 1\}.$$

Then we have

$$C_f = K \cup \left((A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup \ldots \cup (A_{n-m-1} \setminus A_{n-m}) \cup A_{n-m}\right).$$

We will divide the proof into three steps. First we will see that one can always assume that $K = \emptyset$, that is, $C_f = \{Df = 0\}$. Then, in the second step, we will prove that $\mathcal{L}^m(f(A_{n-m})) = 0$. This will conclude the proof of the theorem in

---

1 Indeed we will see that, by our proof, one also has the following result: if $f \in W^{1,p}_{\text{loc}}(\Omega, \mathbb{R}^n)$, with $p > n$, and $E$ is a $\mathcal{L}^n$-null set, then $\mathcal{L}^m(f(E)) = 0$ (see the first part in the proof of Step 2). This fact, the classical area formula for Lipschitz functions, and a standard approximation of $W^{1,p}$ functions with Lipschitz ones, imply the validity of the area formula also in the Sobolev case.
the case \( n = m + 1 \) (as, in this case, \( C_f = K \cup A_1 \)) and will allow us to start an induction argument on \( n - m \). In fact, once we have proved the second step, we can assume that the theorem holds for \( W^{n-m,k} \) maps from an open subset of \( \mathbb{R}^{n-1} \) to \( \mathbb{R}^m \). In the third step, thanks to an Implicit Function Theorem, we will reduce the dimension from \( n \) to \( n - 1 \) and we will conclude the proof by the inductive hypothesis.

**Step 1 (we can assume \( K = \emptyset \)).** This is essentially Step 1 in the proof of Theorem 4.1 in [3].

Let \( K_i := \{ x \in \Omega \mid \text{rank} \, D f(x) = i \} \), \( 1 \leq i \leq m - 1 \), and fix \( \varpi \in K_i \). We can assume det \( \left( \frac{\partial (f_1, \ldots, f_i)}{\partial (x_1, \ldots, x_i)} \right) (\varpi) \neq 0 \). Then, in a small relatively compact neighborhood \( V \) of \( \varpi \), we can take as coordinates \((y_1, \ldots, y_n) = Y(x) := (f_1(x), \ldots, f_i(x), x_{i+1}, \ldots, x_n)\). So, defining \( X := Y^{-1}, \) \( f \) assumes the form

\[
f(X(y)) = (y_1, \ldots, y_i, g(y_1, \ldots, y_n)).
\]

Now, since \( Y \in W^{k,p}(V), \) \( D Y \) is invertible, \( k \geq 2 \) and \( p > n \), it is simple to verify that \( X \) is still \( W^{k,p} \). Moreover, for \( k \geq 2 \) and \( p > n \) the composition of functions in \( W^{k,p}_{loc} \) is still \( W^{k,p}_{loc} \), and so we deduce that \( g \in W^{k,p}(\tilde{V}, \mathbb{R}^{m-i}), \) with \( \tilde{V} := Y(V) \). In these coordinates we have

\[
D(f \circ X)(y) = \begin{pmatrix} I_{i, i} & 0 \\ 0 & D(g|_{Y_1, \ldots, Y_i}) \end{pmatrix}
\]

where \( g|_{(y_1, \ldots, y_i)} : \tilde{V}|_{(y_1, \ldots, y_i)} \to \mathbb{R}^{m-i} \) is defined by \( g|_{(y_1, \ldots, y_i)}(y_{i+1}, \ldots, y_n) = g(y_1, \ldots, y_n), \) with \( \tilde{V}|_{(y_1, \ldots, y_i)} := \{(z_1, \ldots, z_{n-i}) \in \mathbb{R}^{n-i} \mid (y_1, \ldots, y_i, z_1, \ldots, z_{n-i}) \in \tilde{V}\} \).

Observing that rank \( D(f \circ X) = \text{rank} \, D f = i \), thanks to the Slicing Theorem for Sobolev functions (see [4] paragraph 4.9.2) we have, for \( \mathcal{L}^{i} \)-a.e. \( (y_1, \ldots, y_i), \)

\[
g|_{(y_1, \ldots, y_i)} \in W^{k,p}(\tilde{V}|_{(y_1, \ldots, y_i)}, \mathbb{R}^{m-i})
\]

and

\[
D(g|_{(y_1, \ldots, y_i)}) = 0 \quad \text{on} \quad \tilde{V}|_{(y_1, \ldots, y_i)} \cap Y(K_i).
\]

Once we will have proved the result in the case \( K = \emptyset \), applying it to \( g|_{(y_1, \ldots, y_i)} \) we get

\[
\mathcal{L}^{m-i}(g|_{(y_1, \ldots, y_i)}(\tilde{V}|_{(y_1, \ldots, y_i)} \cap Y(K_i))) = 0 \quad \text{for} \quad \mathcal{L}^{i} \text{-a.e.} \quad (y_1, \ldots, y_i).
\]

By Fubini’s Theorem, \( \mathcal{L}^{m}(f(V \cap K_i)) = \mathcal{L}^{m}(f \circ X(\tilde{V} \cap Y(K_i))) = 0 \), and this concludes the proof of the reduction to the case \( K = \emptyset \).

**Step 2 (\( \mathcal{L}^{m}(f(A_{n-m})) = 0 \)).** We recall that, by the remark made at the beginning of the proof, we can assume that \( \Omega \) is bounded; this implies, in particular, that \( A_{n-m} \) has finite Lebesgue measure.

Let \( x \in A_{n-m} = A_{k-1}, \) \( y \in B(x, r) \) with \( r \) such that \( B(x, r) \subset \Omega \). As \( D^{k-1} f \in W^{1,p}(\Omega) \), by the Taylor formula with integral remainder and by (1) we get

\[
|f(y) - f(x)| \leq \int_0^1 \frac{(1-t)^{k-2}}{(k-2)!} |D^{k-1} f(x + t(y-x)) - D^{k-1} f(x)||y-x|^{k-1} \, dt
\]

\[
\leq C r^{\frac{k}{p}} \left( \int_{B(x, r)} |D^{k} f(z)|^p \, dz \right)^{\frac{1}{p}},
\]
which implies
\[ |f(y) - f(x)|^m \leq C \left( \int_{B(x,r)} |D^k f(z)|^p \, dz \right)^{\frac{m}{p}}. \]

Now, using Young’s inequality with exponents \( \frac{p}{p-m} \) and \( \frac{p}{m} \) and taking \( r = |y-x| \), we get
\[ |f(y) - f(x)|^m \leq C|x-y|^{\frac{m}{p-m}} + C \left( \int_{B(x,|y-x|)} |D^k f(z)|^p \, dz \right). \]

As \( k = n - m + 1 \) and \( m(k-1) \geq k-1 \), we have \( km \geq k + m - 1 = n \), which implies \( \frac{pm}{p-m}(k-\frac{1}{p}) = \frac{m}{p-m}(pk-n) \geq n \). So, for \( |y-x| \leq r \leq 1 \) and \( x \in A_{k-1} \), we have the estimate
\[ |f(y) - f(x)|^m \leq C \int_{B(x,r)} (1 + |D^k f(z)|^p) \, dz. \]

We now write \( A_{n-m} = F_1 \cup F_2 \), where
\[ F_1 := \{ \text{density points for } A_{n-m} \} \cap \{ \text{Lebesgue points of } |D^k f|^p \} \]
and
\[ F_2 := A_{n-m} \setminus F_1. \]

It’s a standard result in measure theory that \( \mathcal{L}^m(f(F_2)) = 0 \). Let us now show that
\[ \mathcal{L}^m(f(F_2)) = 0. \]

Fix \( \varepsilon > 0 \) small. \( \mathcal{L}^n(F_2) \) being 0, there exists an open set \( E_{\varepsilon} \supset F_2 \) such that \( E_{\varepsilon} \subset \Omega \) and \( \mathcal{L}^n(E_{\varepsilon}) \leq \varepsilon \). For any \( x \in F_2 \) we take a ball \( B_x = B(x, r_x) \) such that \( B_x \subset E_{\varepsilon} \). We now define \( \rho_x := \text{diam } f(B_x) \), and we consider the covering of \( f(F_2) \) given by \( F = \{ B(f(x), \rho_x) \}_{x \in F_2} \). By Vitali’s Covering Theorem (see [4, Paragraph 1.5.1]), there exists \( G = \{ B(f(x_i), \rho_{x_i}) \}_{i \in I} \) a finite or countable collection of disjoint balls in \( F \) such that
\[ F_2 \subset \bigcup_{i \in I} B(f(x_i), 5\rho_{x_i}). \]

By the definition of \( \rho_{x_i} \), we have
\[ f(B_{z_i}) \subset B(f(x_i), \rho_{x_i}), \]
which implies that the balls \( B_{z_i} \) are also disjoint. Therefore, by (3) we get
\[ \mathcal{L}^m(f(F_2)) \leq 5^m \sum_{i \in I} \mathcal{L}^m(B(f(x_i), \rho_{x_i})) = C_m \sum_{i \in I} (\text{diam } f(B_{z_i}))^m \leq C \sum_{i \in I} \int_{B_{z_i}} (1 + |D^k f(z)|^p) \, dz \leq C \int_{E_{\varepsilon}} (1 + |D^k f(z)|^p) \, dz, \]
where \( C_m = 5^m \mathcal{L}^m(B(0, 1)) \). Letting \( \varepsilon \to 0 \), since \( \mathcal{L}^n(E_{\varepsilon}) \leq \varepsilon \) we obtain
\[ \mathcal{L}^m(f(F_2)) = 0 \] as wanted.

In order to prove that \( \mathcal{L}^m(f(F)) = 0 \), we have to show that \( \mathcal{L}^m(f(F_1)) = 0 \). As we do not have that \( \mathcal{L}^m(F_1) = 0 \), we see that the inequality (3) does not suffice, but in this case, as \( F_1 \) consists of the density points of \( F \), we will get a better estimate for \( |f(y) - f(x)| \) when \( x, y \in F_1 \).
Fix $P \in \mathbb{N}$ large. For any $x \in F_1$ there exists $r_x > 0$ small such that $B(x, 2r_x) \subset \Omega$ and the following hold:

$$\frac{\mathcal{L}^n(B(x, r_x) \cap F_1)}{\mathcal{L}^n(B(x, r_x))} = \frac{\mathcal{L}^n(B(x, r_x) \cap A_{n-m})}{\mathcal{L}^n(B(x, r_x))} \geq 1 - \frac{1}{2(2P)^n},$$

$$\int_{B(x, 2r_x)} (1 + |D^k f(z)|^p) \, dz \leq 2 \left(1 + |D^k f(x)|^p\right)$$
and

$$\frac{1}{2} \left(1 + |D^k f(x)|^p\right) \leq \int_{B(x, r_x) \cap F_1} \left(1 + |D^k f(z)|^p\right) \, dz$$

(this can always be done since $x$ is both a Lebesgue point of the integrated function and a density point of $F_1$). These equations imply a sort of doubling property: if $x \in F_1$, then

$$\int_{B(x, 2r_x)} \left(1 + |D^k f(z)|^p\right) \, dz \leq 2^{n+1} \mathcal{L}^n(B(x, r_x)) \left(1 + |D^k f(x)|^p\right)$$

$$\leq 2^{n+2} \mathcal{L}^n(B(x, r_x) \cap F_1) \left(1 + |D^k f(x)|^p\right)$$

$$\leq 2^{n+3} \int_{B(x, r_x) \cap F_1} \left(1 + |D^k f(z)|^p\right) \, dz.$$

Moreover, for each $y \in F_1 \cap B(x, r_x)$, there exist $P + 1$ points $\{x_0, \ldots, x_P\} \subset F_1$, with $x_0 = y$ and $x_P = x$, such that

$$|x_i - x_{i-1}| \leq \frac{2r_x}{P} \quad \forall 1 \leq i \leq P.$$ 

Indeed, first take $y_1, \ldots, y_{P-1}$ to be $P - 1$ points on the line segment $[y, x]$ such that $|y_i - y_{i-1}| = \frac{|y - x|}{P}$ and then observe that, by (4), $B(y_i, \frac{2r_x}{P}) \cap F_1$ is not empty for each $i$, and so it suffices to take a point $x_i$ in that set. By this and (2), it follows that

$$|f(y) - f(x)| \leq \sum_{i=1}^{P} |f(x_i) - f(x_{i-1})|$$

$$\leq C \sum_{i=1}^{P} |x_i - x_{i-1}|^{k - \frac{n}{p}} \left(\int_{B(x, \frac{2r_x}{P})} |D^k f(z)|^p \, dz\right)^{\frac{1}{p}}$$

$$\leq C \sum_{i=1}^{P} \left(\frac{2r_x}{P}\right)^{k - \frac{n}{p}} \left(\int_{B(x, 2r_x)} |D^k f(z)|^p \, dz\right)^{\frac{1}{p}}$$

whenever $y \in B(x, r_x)$. Again using Young’s inequality, we get

$$|f(y) - f(x)|^m \leq CP^m(1 - k + \frac{p}{n}) \int_{B(x, 2r_x)} \left(1 + |D^k f(z)|^p\right) \, dz \quad \forall y \in B(x, r_x).$$

Thus by (5) we obtain that, for all $x \in F_1$,

$$|f(y) - f(x)|^m \leq CP^m(1 - k + \frac{p}{n}) \int_{B(x, r_x) \cap F_1} \left(1 + |D^k f(z)|^p\right) \, dz \quad \forall y \in B(x, r_x).$$

We are now able to prove that $\mathcal{L}^m(f(F_1)) = 0$.

For any $x \in F_1$ we take the ball $B_x = B(x, r_x)$, where $r_x$ was defined above. We now define $\rho_x := \text{diam } (B_x \cap F_1)$, and we consider the covering of $f(F_1)$
given by \( F = \{ B(f(x), \rho_x) \}_{x \in F_1} \). Using again Vitaly’s theorem we find \( G = \{ B(f(x_i), \rho_{x_i}) \}_{i \in I} \) a finite or countable collection of disjoint balls in \( F \) such that
\[
F_1 \subset \bigcup_{i \in I} B(f(x_i), 5\rho_{x_i}).
\]
In this case, by the definition of \( \rho_{x_i} \), we have
\[
f(B_{x_i} \cap F_1) \subset B(f(x_i), \rho_{x_i}),
\]
which implies that the sets \( B_{x_i} \cap F_1 \) are disjoint. Arguing as for \( F_2 \), thanks to \( \text{(ii)} \) we obtain
\[
\mathcal{L}^m(f(F_1)) \leq C \sum_{i \in I} (\text{diam } B(x_i, \cap F_1))^m
\leq C P^m(1-k+p) \sum_{i \in I} \int_{B_{x_i} \cap F_1} (1 + |D^k f(z)|^p) \, dz
\leq C P^m(1-k+p) \int_\Omega (1 + |D^k f(z)|^p) \, dz,
\]
and we conclude by letting \( P \to +\infty \), as \( k \geq 2 > 1 + \frac{n}{p} \).

**Step 3** \( \mathcal{L}^m(f(A_{s-1} \setminus A_s)) = 0 \), for \( 2 \leq s \leq k - 1 \). Fix \( \varpi \in A_{s-1} \setminus A_s \). In order to prove the claim, it suffices to show that there exists an open neighborhood \( V \) of \( \varpi \) such that \( \mathcal{L}^m(f((A_{s-1} \setminus A_s) \cap V)) = 0 \). We recall that, by what we already said, our function is \( C^{k-1, \alpha(p)} \). Now, as \( \varpi \in A_{s-1} \), \( f \) is \( (s-1) \)-flat at \( \varpi \), but some partial derivative of order \( s \) is not zero. Hence we may assume that
\[
\partial_n w(\varpi) \neq 0, \quad w(\varpi) = \partial_1 \ldots \partial_{s-1} f(\varpi) = 0.
\]
We observe that \( w \in C^{k-s, \alpha(p)} \), and hence, by the Implicit Function Theorem, there is an open neighborhood \( V \) of \( \varpi \) such that \( V \cap \{ w = 0 \} \) is an \((n-1)\)-dimensional \( C^{k-s, \alpha(p)} \)-graph, and so we have \( V \cap A_{s-1} \subset g(W) \), where \( W \subset \mathbb{R}^{n-1} \) is open and \( g : W \to \mathbb{R}^n \) is \( C^{k-s, \alpha(p)} \).

Let us now consider the subset \( A^* \subset W \) defined by \( A^* := \{ x \in W \mid g(x) \in A_{s-1} \} \). By Theorem \( \text{(ii)} \) there exists a function \( F : W \to \mathbb{R}^m \) of class \( C^{k-1} \) such that:

(i) \( F(x) = f(g(x)) \) for any \( x \in A^* \);
(ii) \( DF(x) = 0 \) for any \( x \in A^* \).

Therefore we have \( f(A_{s-1} \cap V) \subset F(C_F \cap W) \), where \( C_F \) denotes the critical set of \( F \). So it suffices to prove that \( \mathcal{L}^m(C_F \cap W) = 0 \), and this follows by the induction hypothesis since
\[
F \in C^{k-1}(W, \mathbb{R}^m) \hookrightarrow W^{k-1,p}_{loc}(W, \mathbb{R}^m).
\]

\( \square \)

**Acknowledgements**

It’s a pleasure to thank Albert Fathi for useful discussions and to gratefully acknowledge the hospitality of the École Normale Supérieure of Lyon, where this paper was written. I also thank David Preiss and the referee of the paper for their helpful comments.
References


Université de Nice Sophia-Antipolis, Laboratoire J.-A. Dieudonné, CNRS UMR 6621, Parc Valrose, 06108 Nice Cedex 02, France

E-mail address: figalli@unice.fr