Zeitschrift für Analysis und ihre Anwendungen
Journal of Analysis and its Applications
Volume 34 (2015), 321–342
DOI: 10.4171/ZAA/1542

The Weak Inverse Mapping Theorem

Daniel Campbell, Stanislav Hencl and František Konopecký

Abstract. We prove that if a bilipschitz mapping f is in $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n,\mathbb{R}^n)$ then the inverse f^{-1} is also a $W^{m,p}_{\mathrm{loc}}$ class mapping. Further we prove that the class of bilipschitz mappings belonging to $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n,\mathbb{R}^n)$ is closed with respect to composition and multiplication without any restrictions on $m, p \geq 1$. These results can be easily extended to smooth n-dimensional Riemannian manifolds and further we prove a form of the implicit function theorem for Sobolev mappings.

Keywords. Bilipschitz mappings, inverse mapping theorem

Mathematics Subject Classification (2010). Primary 46E35, secondary 26B10

1. Introduction

It is well known that for a mapping of $\mathbb{R}^n$ to $\mathbb{R}^n$, which is of class $\mathcal{C}^m$ and has positive Jacobian at some point x we can find neighbourhoods of x and $f(x)$ such that the restriction of f is a $\mathcal{C}^m$ class diffeomorphism of the two neighbourhoods. Building on [7] and [9], we ask if this result can be extended to classes of Sobolev mappings.

Since the seminal paper of Arnold [2] a variety of techniques have been applied to hydrodynamics and other partial differential equations using certain properties of the spaces of Sobolev diffeomorphisms on smooth Riemannian manifolds. We refer the reader to [9] for more detailed motivation and applications.

In [9] the authors considered the regularity of the inverse of a Sobolev diffeomorphism and the composition of a Sobolev mapping in $W^{s+r,2}(\mathbb{R}^n,\mathbb{R}^n)$ with a mapping pertaining to a certain class of $W^{s,2}_{\mathrm{loc}}$ Sobolev $\mathcal{C}^1$ diffeomorphisms, $s \in \mathbb{R}$. Our result is a generalization of the above, in that we remove the condition $s > \frac{n}{2} + 1$ completely and consider $p \in [1,\infty]$ arbitrary. We also relax

D. Campbell: Department of Mathematics, University of Hradec Králové, Rokitanského 62, 500 03 Hradec Králové, Czech Republic; daniel.campbell@uhk.cz
S. Hencl, F. Konopecký: Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00 Prague 8, Czech Republic;
hencl@karlin.mff.cuni.cz; frakon@gmail.com

the condition that f needs to be a diffeomorphism and require only bilipschitz regularity. Our result is also an extension of that proved in [7, Theorem 1.3], where it was assumed that $m = 2$. Our main result is as follows.

Theorem 1.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and $f : \Omega \to \Omega' = f(\Omega)$ be such that*

$$f \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Then

$$f^{-1} \in W^{m,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^n).$$

Further, if $D^m f \in \mathrm{BV}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n^{m+1}})$ then $D^m f^{-1} \in \mathrm{BV}_{\mathrm{loc}}(\Omega', \mathbb{R}^{n^{m+1}})$.

The counterexamples given in [7] (see also Remark 3.5) show that the bilipschitz condition is vital to guarantee that the inverse is Sobolev or BV, i.e. the result may fail if f is not Lipschitz and it may fail if f^{-1} is not Lipschitz. For the definition of bilipschitz mappings, $\mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, see the preliminaries. It is not difficult to check that the Sobolev imbedding theorem gives that if $r \in \mathbb{N}$ is such that $r < m - \frac{n}{p}$ then f and f^{-1} are $\mathcal{C}^r$ mappings.

In [9] the authors pose the question whether the composition of two diffeomorphisms in $W^{s,2}(M, M)$, $f \circ g$, is a $W^{s,2}(M, M)$ map, $s \in \mathbb{N}, s > \frac{n}{2}$. We prove this is true for $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ mappings if the interior function is bilipschitz and the exterior Lipschitz. Our result places no constraint on m, p other than $m \in \mathbb{N}$ and $p \in [1, \infty]$. Note that previous results about the regularity of the composition (see e.g. [3] and [12]) usually assume that all lower order derivatives of f are bounded which is not necessary if we moreover assume that g is bilipschitz. The results from [9] are extended also for fractional order Sobolev spaces. We have not pursued this direction but we don't see any obstacles in doing so. Our result is as follows.

Theorem 1.2. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and $g : \Omega \to \Omega' = f(\Omega)$ be such that*

$$g \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n), \quad f \in W^{m,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^n) \cap W^{1,\infty}(\Omega', \mathbb{R}^n).$$

Then

$$f \circ g \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Again this result may fail if g^{-1} is not Lipschitz and it may fail if f is not Lipschitz even for bilipschitz g - see Remark 3.4. We also prove the following result about the product of two functions in the class considered above.

Theorem 1.3. *Let $\Omega, \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and*

$$f, g \in W^{m,p}_{\mathrm{loc}}(\Omega) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega).$$

Then

$$fg \in W^{m,p}_{\mathrm{loc}}(\Omega) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega).$$

It is not difficult to check that the Sobolev imbedding theorem gives that if $r \in \mathbb{N}$ is such that $r < m - \frac{n}{p}$ then $g \circ f$ and fg are $\mathcal{C}^r$ mappings.

Let us give the brief idea of our proof of the main result Theorem 1.1. By twice differentiation $D^2(f \circ f^{-1}) = D^2(id) = 0$ we obtain the identity

$$D^2 f^{-1}(y) = -Df^{-1}(y) D^2 f(f^{-1}(y)) Df^{-1}(y) Df^{-1}(y),$$

see Preliminaries for the interpretation of the higher order derivatives. Using the Leibniz rule and the chain rule we derive this identity further and we express $D^k f^{-1}$ as a product of lower order derivatives of f and f^{-1}. We estimate the integrability of the lower order terms using induction and the famous Gagliardo-Nirenberg interpolation inequality. Simple use of Hölder's inequality gives our final claim. Let us note that the simple use of the Sobolev embedding theorem would not be sufficient to prove the claim without some extra assumption on the lower order derivatives. Fortunately the Gagliardo-Nirenberg interpolation inequality gives us a better integrability of the lower order terms which gives us exactly the desired result.

It is not difficult to show that similar results hold on smooth n-dimensional Riemannian manifolds. In Section 4 we show that it is enough to apply our Euclidean result for the composition with reference maps. Finally in Section 5 we prove a variant of the **implicit mapping theorem** for Sobolev mappings using our inverse mapping theorem 1.1.

2. Preliminaries

2.1. Results on Sobolev functions. We start by defining locally bilipschitz mappings.

Definition 2.1. Let $\Omega \subset \mathbb{R}^n$ be open and $u : \Omega \to \mathbb{R}^d$. The space $\mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ is the class of mappings $u : \Omega \to \mathbb{R}^d$ such that for all $x_0 \in \Omega$ there exists some $\delta > 0$ and $C_1, C_2 > 0$, such that for all $x, x' \in \Omega \cap B(x_0, \delta)$ it holds that

$$C_1|x - x'| < |u(x) - u(x')| < C_2|x - x'|.$$

For the following Theorem see [1, Theorem 3.16 and Corollary 3.19]:

Theorem 2.2. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $u : \Omega \to \mathbb{R}^d$. Suppose that $F : \Omega \to \Omega' = F(\Omega)$ is Lipschitz and a homeomorphism.*

1. *If $u \in BV_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ then $u \circ F^{-1} \in BV_{\mathrm{loc}}(\Omega', \mathbb{R}^d)$.*
2. *If $u \in W^{1,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ and F^{-1} is Lipschitz, then $u \circ F^{-1} \in W^{1,1}_{\mathrm{loc}}(\Omega', \mathbb{R}^d)$ and*

$$Du \circ F^{-1}(y) = Du(F^{-1}(y)) DF^{-1}(y) \quad \textit{for almost all } y \in \Omega'.$$

We will refer to the following lemma as the product rule or the Leibniz rule.

Lemma 2.3. *Let $\Omega \subset \mathbb{R}^n$ be open and $f, g \in W^{1,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$. Suppose that*

$$fDg, gDf \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^d).$$

Then

$$D(f(x)g(x)) = g(x)Df(x) + f(x)Dg(x) \quad \textit{for almost all } x \in \Omega.$$

Similarly we can prove the formula for differentiation of three or more terms under the assumption that all summands are integrable.

Proof. By considering the component functions we may assume that $d = 1$. We approximate the functions by truncation. Our truncated functions belong to $W^{1,1}_{\mathrm{loc}}(\Omega) \cap L^\infty(\Omega)$ and here we can use the standard convolution approximations to get our result for the truncated functions. The function $C(|gDf| + |fDg|)$ is integrable and dominates the derivative of our truncation approximations. This however gives us the desired equality almost everywhere. $\square$

The following theorem [7, Theorem 1.3] will be a start for our induction process.

Theorem 2.4. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open, $p \geq 1$ and suppose that $f : \Omega \to \Omega'$ is a bilipschitz mapping. If $Df \in W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n^2})$, then $Df^{-1} \in W^{1,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^{n^2})$.*

The Sobolev Embedding Theorem is well known.

Theorem 2.5. *Let $\Omega \subset \mathbb{R}^n$ be open and have a Lipschitz boundary. Further let $p \in [1, n)$ and $f \in W^{1,p}(\Omega)$. Then $f \in L^{\frac{np}{n-p}}(\Omega)$ and moreover*

$$\|f\|_{L^{\frac{np}{n-p}}(\Omega)} \leq c\|f\|_{W^{1,p}(\Omega)}.$$

Further if $f \in W^{k,p}(\Omega)$ and $kp < n$ then for every $i \in \{0,, \ldots, k-1\}$ we have $f \in W^{i,p_i}(\Omega)$ where $p_i = \frac{np}{n-(k-i)p}$. If $f \in W^{1,p}(\Omega)$ for some $p > n$ then $f \in L^\infty(\Omega)$. This also holds for mappings with values in $\mathbb{R}^d$.

The following theorem is referred to as the Gagliardo-Nirenberg interpolation inequality and is a result of [13].

Theorem 2.6. *Let $u : \mathbb{R}^n \to \mathbb{R}$, $q, r \in [1, \infty]$ and $\hat{k} \in \mathbb{N}$. Further let $j \in \{1, \ldots, \hat{k}\}$, $\hat{p} \in [1, \infty)$ and $\alpha \in [\frac{j}{\hat{k}}, 1]$ be such that*

$$\frac{1}{\hat{p}} = \frac{j}{n} + \alpha\Big(\frac{1}{r} - \frac{\hat{k}}{n}\Big) + \frac{1-\alpha}{q}.$$

Then $u \in L^q(\mathbb{R}^n)$ and $|D^{\hat{k}}u| \in L^r(\mathbb{R}^n)$ implies that $|D^j u| \in L^{\hat{p}}(\mathbb{R}^n)$ and this embedding is continuous.

Let us now prove that this result can be extended as follows.

Theorem 2.7. *Let $u : \mathbb{R}^n \to \mathbb{R}$ and $q \in [1,\infty]$ and $\hat{k} \in \mathbb{N}$. Further let $j \in \{1,\dots,\hat{k}-1\}$, $\hat{p} \in [1,\infty)$ and $\alpha \in [\frac{j}{\hat{k}},1]$ be such that*

$$\frac{1}{\hat{p}} = \frac{j}{n} + \alpha\Big(1 - \frac{\hat{k}}{n}\Big) + \frac{1-\alpha}{q}.$$

Then $u \in L^q(\mathbb{R}^n)$ and $D^{\hat{k}-1}u \in \mathrm{BV}(\mathbb{R}^n)$ implies that $|D^j u| \in L^{\hat{p}}(\mathbb{R}^n)$.

Sketch of the proof. For f_k, the convolution approximations of f, it holds that $f_k \in W^{m,1}(\mathbb{R}^n)$ and we may apply Theorem 2.6. Since $||f_k||_q \leq ||f||_q$ and $||D^m f_k||_1 \leq ||D^m f||_1$ (where $||D^m f||_1$ signifies the total variation) we have that $D^j f_k$ is a bounded sequence and by moving if necessary to a subsequence we find for every $(a_1,\dots,a_j) \in \{1,\dots,n\}^j$ a $w_{a_j,\dots,a_1} \in L^{\hat{p}_j}(\mathbb{R}^n)$ such that for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} D_{a_j,\dots,a_1} f_k \varphi \to \int_{\mathbb{R}^n} w_{a_j,\dots,a_1}\varphi$$

and clearly

$$\int_{\mathbb{R}^n} D_{a_j,\dots,a_1} f_k \varphi = (-1)^j \int_{\mathbb{R}^n} f_k D_{a_j,\dots,a_1}\varphi \to (-1)^j \int_{\mathbb{R}^n} f D_{a_j,\dots,a_1}\varphi.$$

This however implies that $D_{a_j,\dots,a_1} f = w_{a_j,\dots,a_1} \in L^{\hat{p}_j}(\mathbb{R}^n)$. □

We will use the so called ACL classification of Sobolev mappings.

Theorem 2.8. *Let $\Omega \subset \mathbb{R}^n$ be open and let $f \in W^{1,p}(\Omega)$ then there exists a representative $\hat{f}$ of f such that $\hat{f}$ is absolutely continuous on almost all lines parallel to each of the coordinate axes. Further the classical partial derivatives of $\hat{f}$ equal the weak derivatives of f almost everywhere.*

It is not difficult to use the previous theorem recurrently to see that if $f \in W^{k,p}(\Omega)$ then there exists a representative of f such that for all $1 \leq j \leq k-1$, the derivative $D^j f$ is absolutely continuous on almost all lines parallel to the coordinate axis. Further the classical partial derivatives of our representative equal the weak derivatives of f almost everywhere.

2.2. Representation of higher order derivatives. During the course of this work we will need to represent and work with higher order derivatives. We refer the reader to [11] for more information about multi-linear algebra which could be used to represent our higher-order derivatives.

To shorten our notation D_i, $i \in \{1, 2, \dots, n\}$, will denote the weak derivative in the direction of the i-th canonic basis vector. Given a finite sequence $a_1, a_2, \dots, a_m \in \{1, 2, \dots, n\}$ we define the symbol

$$D_{a_m,\dots,a_2,a_1} f(x) = D_{a_m}\Big(\cdots \big(D_{a_2}\big(D_{a_1} f(x)\big)\big)\cdots\Big).$$

We will use the following notation for the components of a mapping $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x) = (f^1(x), f^2(x), \dots, f^n(x))$. This notation will be useful as in the chain rule we will always sum over indices, one of which is a superscript and the other a subscript. This corresponds to common practice in tensor notation.

Given some $f \in W^{m,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ we can define the symbol $D^m f$ as the mapping from $\{1, 2, \dots, n\}^m \to L^1(\Omega, \mathbb{R}^n)$ such that technically speaking for $a_i \in \{1, 2, \dots, n\}$ we have

$$D^m f(a_1, \dots, a_m) = D_{a_m,\dots,a_1} f = D_{a_m}(\cdots D_{a_2}(D_{a_1} f))\cdots),$$

but we will write

$$D^m f = \Big(D_{a_m,\dots,a_1} f\Big)_{a_1,\dots,a_m \in \{1,2,\dots,n\}}.$$

It therefore follows that for $f \in W^{m,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ we can identify $D^m f$ with an element of the set $L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^{n^{m+1}})$. It is an easy result of the definition of the weak derivative that the order of partial derivatives is interchangeable, i.e.

$$D_{a_m,\dots,a_1} f = D_{a_{\pi(m)},\dots,a_{\pi(1)}} f$$

for any $\pi \in S_m$ the symmetric group. It suffices to take the definition of the weak derivatives, where in the integral we can change the order of derivatives on the test function as it is smooth. Thus our derivatives $D^j f$ are symmetric.

We shall now expound the iterated chain rule, which is a result of repeatedly using the chain rule and the product rule, for a pair of smooth mappings f and g. We will later prove that the same holds for Sobolev mappings given g is bilipschitz.

Let f and g be smooth. Clearly

$$D(f \circ g)(x) = Df(g(x))Dg(x)$$

where the multiplication above is standard matrix multiplication. Now apply D again and use the product rule. We get

$$D_{j,i}(f \circ g)(x) = \sum_{k,l=1}^{n} D_{k,l} f(g(x))\big(D_i g(x)\big)^k \big(D_j g(x)\big)^l + \sum_{l=1}^{n} D_l f(x)\big(D_{j,i} g(x)\big)^l.$$

This can be symbolically rewritten as

$$D^2(f \circ g)(x) = D^2 f(g(x))Dg(x)Dg(x) + Df(x)D^2 g(x), \tag{1}$$

where we expand the concept of matrix multiplication as follows: in the first term we sum over $\{1,2,\dots,n\}^2$, all second partials of f and the components of the two Dg terms and over $\{1,2,\dots,n\}$, the first partials of f and the components of D^2g in the second term.

Notice that if we make the following definition of $\mathcal{K}_m$ with $m \in \mathbb{N}$

$$\mathcal{K}_m = \Big\{(k_0,\dots,k_m) \in \{0,1,\dots,m\}^{m+1} : \\ (k_0 \geq 1) \ \&\ (k_i = 0 \Leftrightarrow i > k_0) \ \&\ \Big(\sum_{i=1}^{m} k_i = m\Big)\Big\}$$

we may write the above as follows

$$D^2(f \circ g)(x) = \sum_{k\in\mathcal{K}_2} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x).$$

We will now prove the following equality for all $m \in \mathbb{N}$ by induction

$$D^m\big(f \circ g\big)(x) = \sum_{k\in\mathcal{K}_m, \chi\in X_k} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x), \tag{2}$$

where we define X_k below. Here k_0 corresponds to the order of derivative of f and other k_i correspond to the derivatives of g. Note that the definition of $\mathcal{K}_m$ gives us $\sum_{i=1}^m k_i = m$ since in each step we differentiate some term with g once more and also $k_i = 0, i > k_0$ since after deriving f k_0 times we have at most k_0 terms with g. In fact, for given numbers $\{k_i\}_{i=1}^m$ there is a number of permissible permutations (orderings in which the derivatives may be applied), which we will denote as X_k for some $k \in \mathcal{K}_m$. This corresponds to the fact that each term $D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x)$ is multiplied by some fixed natural number $\#X_k$.

It is no problem for us to work with the identity (2) and with its interpretation. In our proof we will deduce the integrability of each term in the product $D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x)$ and then we will apply the Hölder inequality. Therefore we do not need to care which component is multiplied with which component of the next term. We have finitely many terms and each of them can be estimated in the same way.

For our component-wise notation we will need to take a finite sequence in $\{1,2,\dots,n\}$ and apply each corresponding derivative to one of our factors. We define

$$X_k = \Big\{\chi : \{1,\dots,m\} \to \{1,\dots,k_0\};\ \#\{j : \chi(j) = i\} = k_i; \\ \min\{j \leq m : \chi(j) = i+1\} > \min\{j \leq m : \chi(j) = i\} : 1 \leq i \leq m-1\Big\}.$$

We carry on to define b^i_j by denoting

$$b^i_j = a_l,$$

where l is the j-th index which is mapped to i by χ, i.e. $\chi(l) = i$ and $\#\{\chi(o) = i : o < l\} = j - 1$. Assume that the following holds for all $m < \hat{m}$

$$D_{a_m,\ldots,a_1}(f\circ g)(x) = \sum_{k\in\mathcal{K}_m}\sum_{\chi\in X_k}\sum_{b^0_1,\ldots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\ldots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\ldots,b^i_1} g(x))^{b^0_i}. \quad (3)$$

Let us now differentiate the equation above, where $m = \hat{m}-1$, by applying $D_{a_{\hat{m}}}$. We apply the Leibnitz rule (product rule) repeatedly and when differentiating $D_{b^0_{k_0},\ldots,b^0_1} f(g(x))$ we also use the chain rule and get

$$\begin{aligned} &D_{a_{\hat{m}},\ldots,a_1}(f\circ g)(x) \\ &= \sum_{\hat{j}=1}^{n}\sum_{k\in\mathcal{K}_{\hat{m}}}\sum_{\chi\in X_k}\sum_{b^0_1,\ldots,b^0_{k_0}=1}^{n} D_{\hat{j},b^0_{k_0},\ldots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\ldots,b^i_1} g(x))^{b^0_i} \cdot (D_{a_{\hat{m}}} g(x))^{\hat{j}} \\ &\quad + \sum_{\hat{i}=1}^{k_0}\sum_{k\in\mathcal{K}_{\hat{m}}}\sum_{\chi\in X_k}\sum_{b^0_1,\ldots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\ldots,b^0_1} f(g(x)) \\ &\qquad\qquad \times \prod_{i\in\{1,\ldots k_0\}\setminus\{\hat{i}\}} (D_{b^i_{k_i},\ldots,b^i_1} g(x))^{b^0_i} \cdot (D_{a_{\hat{m}} b^{\hat{i}}_{k_{\hat{i}}},\ldots,b^{\hat{i}}_1} g(x))^{b^0_{\hat{i}}}. \end{aligned} \quad (4)$$

We can express this equation using the same notation as in (3). To prove this consider two terms in the above sums, firstly where $k_0 = s$, secondly where $k_0 = s+1$ and add the first line of (4) for $k_0 = s$ and the second line of (4) for $k_0 = s+1$. Putting $M = \{k \in \mathcal{K}_{\hat{m}} : k_0 = s+1\}$ we get

$$\sum_{k\in M}\sum_{\chi\in X_k}\sum_{b^0_1,\ldots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\ldots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\ldots,b^i_1} g(x))^{b^0_i}.$$

Summing this over s gives our desired result.

It will be very useful to be able to express the above in some more concise way, therefore we will introduce the following convention. We will shorten

$$D_{a_m,\ldots,a_1}(f\circ g)(x) = \sum_{k\in\mathcal{K}_m}\sum_{\chi\in X_k}\sum_{b^0_1,\ldots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\ldots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\ldots,b^i_1} g(x))^{b^0_i} \quad (5)$$

by realizing that the order of the derivative of f (i.e. k_0) is the same as the number of Dg factors and we sum over the derivative indices of f and component indices of Dg-type terms as above, writing only

$$D^m f \circ g(x) = \sum_{k \in \mathcal{K}_m, \chi \in X_k} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x) \tag{6}$$

where the previous equalities hold at all points x where $f \circ g$ is defined. As previously mentioned we will extend this result for bilipschitz Sobolev mappings in Sections 3 and 4.

3. Regularity of the inverse

Recall that if we have some mapping $h : \mathbb{R}^n \to \mathbb{R}^l$, we denote the components of h as $h(x) = (h^1(x), h^2(x), \ldots, h^l(x))$.

Lemma 3.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $f : \Omega \to \Omega' = f(\Omega)$ be a bilipschitz mapping such that $f \in W^{2,1}_{\text{loc}}(\Omega, \mathbb{R}^n)$. Then $f^{-1} \in W^{2,1}_{\text{loc}}(\Omega', \mathbb{R}^n)$ and*

$$\begin{aligned} &\Big(D_{j,i} f^{-1}(y)\Big)^k \\ &= -\sum_{l=1}^{n} \sum_{l_1, l_2 = 1}^{n} \Big(D_l f^{-1}(y)\Big)^k \Big(D_{l_2, l_1} f(f^{-1}(y))\Big)^l \Big(D_i f^{-1}(y)\Big)^{l_1} \Big(D_j f^{-1}(y)\Big)^{l_2}, \end{aligned} \tag{7}$$

which may also be written as

$$D^2 f^{-1}(y) = -Df^{-1}(y) D^2 f(f^{-1}(y)) Df^{-1}(y) Df^{-1}(y) \tag{8}$$

for almost all $y \in \Omega'$.

Remark 3.2. In the above lemma, the left-most factor on the right hand side of (8) is the matrix $Df^{-1}(y)$. The reason for this is that it then corresponds to standard composition (multiplication) of matrices. Notice however that in (7) the order of the factors is irrelevant given we sum the correct upper and lower indices.

Proof of Lemma 3.1. First let us note that $f^{-1} \in W^{2,1}_{\text{loc}}(\Omega', \mathbb{R}^n)$ is a direct result of Theorem 2.4. Clearly we have almost everywhere that

$$0 = D(I) = D^2(id) = D^2(f \circ f^{-1}) = D\big(D(f \circ f^{-1})\big)$$

where I is the $(n \times n)$ identity matrix and id the identity mapping on $\mathbb{R}^n$. Now we can use Theorem 2.2 on $f \circ f^{-1}$ as f is bilipschitz. Therefore we get

$$0 = D_j \Bigg(\sum_{l_1 = 1}^{n} D_{l_1} f(f^{-1}(y)) \big(D_i f^{-1}(y)\big)^{l_1} \Bigg)$$

for all $i,j\in\{1,2,\dots,n\}$ and for almost all $y\in\Omega'$. As $Df(f^{-1}), Df^{-1}\in L^\infty(\Omega')$ and f, f^{-1} are of Sobolev type $W^{2,1}_{\text{loc}}$ it is easy to see that we will have no trouble in applying the product rule, Lemma 2.3. Deriving again, using the product rule and then Theorem 2.2 on $Df(f^{-1})$, we get

$$0=\sum_{l_1,l_2=1}^{n} D_{l_2,l_1}f(f^{-1}(y))\big(D_i f^{-1}(y)\big)^{l_1}\big(D_j f^{-1}(y)\big)^{l_2}+\sum_{l=1}^{n} D_l f(f^{-1}(y))\big(D_{j,i}f^{-1}(y)\big)^{l}.$$

Notice that thanks to the Lipschitz qualities of f and f^{-1} we have $Df^{-1}(y)=\big(Df(f^{-1}(y))\big)^{-1}$ almost everywhere in terms of the inverse of matrices. We apply $Df^{-1}(y)$ (matrix multiplication from the left) to get

$$\begin{aligned}0=&\sum_{l=1}^{n}\sum_{l_1,l_2=1}^{n}\Big(D_l f^{-1}(y)\Big)^{k}\Big(D_{l_2,l_1}f(f^{-1}(y))\Big)^{l}\Big(D_i f^{-1}(y)\Big)^{l_1}\Big(D_j f^{-1}(y)\Big)^{l_2}\\&+\Big(D_{j,i}f^{-1}(y)\Big)^{k}\end{aligned}$$

for all components $k\in\{1,2,\dots,n\}$ and for almost all $y\in\Omega'$. □

Lemma 3.3. *Let $\Omega,\Omega'\subset\mathbb{R}^n$ be open. Let $f:\Omega\to\Omega'=f(\Omega)$ be a bilipschitz mapping such that $f\in W^{k,p}_{\text{loc}}(\Omega,\mathbb{R}^n)$. Then*

$$\big|D\big(D^{k-1}f(f^{-1})\big)\big|\in L^p_{\text{loc}}(\Omega')$$

and for all k-tuples $a_1,a_2,\dots a_k\in\{1,2,\dots,n\}$ we have

$$D_{a_k}(D_{a_{k-1},\dots,a_1}f(f^{-1}(y)))=\sum_{l=1}^{n} D_{l,a_{k-1},\dots,a_1}f(f^{-1}(y))\big(D_{a_k}f^{-1}(y)\big)^{l}.$$

Proof. Clearly $D_{a_{k-1},\dots,a_1}f\in W^{1,p}_{\text{loc}}(\Omega,\mathbb{R}^n)$ and therefore belongs also to the Sobolev space $W^{1,1}_{\text{loc}}(\Omega,\mathbb{R}^n)$. Since f^{-1} is bilipschitz, Theorem 2.2 implies that the weak derivative of $D_{a_{k-1},\dots,a_1}f(f^{-1}(\cdot))$ exists and since as f^{-1} is a bilipschitz change of variables we have $D_{a_{k-1},\dots,a_1}f(f^{-1}(\cdot))\in W^{1,p}_{\text{loc}}(\Omega',\mathbb{R}^n)$. From Theorem 2.2 we get

$$D_{a_k}(D_{a_{k-1},\dots,a_1}f(f^{-1}(y)))=\sum_{l=1}^{n} D_l D_{a_{k-1},\dots,a_1}f(f^{-1}(y))(D_{a_k}f^{-1}(y))^{l}\tag{9}$$

for almost all $y\in\Omega'$. Since $|Df^{-1}|\in L^\infty(\Omega')$ we get

$$\big|D\big(D^{k-1}f(f^{-1}(\cdot))\big)\big|\in L^p_{\text{loc}}(\Omega').$$

By applying (9) to all $(a_1,\dots a_{k-1})\in\{1,2,\dots,n\}^{k-1}$ we get that by our convention, described between (5) and (6), that we may write

$$D(D^{k-1}f(f^{-1}(y)))=D^k f(f^{-1}(y))Df^{-1}(y)\quad\text{for almost all } y\in\Omega'.$$ □

Proof of Theorem 1.1. Consider the case $m = 1$. Our claim is trivial as f is bilipschitz. For the case $m = 2$ Theorem 2.4 gives our result. Our proof will continue by induction. We continue to prove the case $m = 3$ explicitly to aid the comprehension of the reader.

Firstly, we know by Lemma 3.1 that (7) holds. We differentiate (7) according to Theorem 2.2, Lemma 3.3 and Lemma 2.3, initially formally, and will afterward verify the assumptions. Thus,

$$\begin{aligned}
&D_{a_3,a_2,a_1} f^{-1}(y)\\
&= -D_{a_3}\Bigg(\sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y)\big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0}\big(D_{a_1} f^{-1}(y)\big)^{l_1}\big(D_{a_2} f^{-1}(y)\big)^{l_2}\Bigg)\\
&= \sum_{l_0,l_1,l_2=1}^{n} D_{a_3,l_0} f^{-1}(y)\big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0}\big(D_{a_1} f^{-1}(y)\big)^{l_1}\big(D_{a_2} f^{-1}(y)\big)^{l_2}\\
&\quad+ \sum_{l_0,l_1,l_2,l_3=1}^{n} D_{l_0} f^{-1}(y)\big(D_{l_3,l_2,l_1} f(f^{-1}(y))\big)^{l_0}\\
&\qquad\qquad \times \big(D_{a_1} f^{-1}(y)\big)^{l_1}\big(D_{a_2} f^{-1}(y)\big)^{l_2}\big(D_{a_3} f^{-1}(y)\big)^{l_3}\\
&\quad+ \sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y)\big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0}\big(D_{a_3,a_1} f^{-1}(y)\big)^{l_1}\big(D_{a_2} f^{-1}(y)\big)^{l_2}\\
&\quad+ \sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y)\big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0}\big(D_{a_1} f^{-1}(y)\big)^{l_1}\big(D_{a_3,a_2} f^{-1}(y)\big)^{l_2},
\end{aligned}$$

which can be summarized, according to our convention, by omitting indices as follows

$$\begin{aligned}
D^3 f^{-1}(y) = {}& D^2 f^{-1}(y) D^2 f(f^{-1}(y)) D f^{-1}(y) D f^{-1}(y)\\
&+ D f^{-1}(y) D^3 f(f^{-1}(y)) D f^{-1}(y) D f^{-1}(y) D f^{-1}(y)\\
&+ D f^{-1}(y) D^2 f(f^{-1}(y)) D^2 f^{-1}(y) D f^{-1}(y)\\
&+ D f^{-1}(y) D^2 f(f^{-1}(y)) D f^{-1}(y) D^2 f^{-1}(y).
\end{aligned}$$

We want to prove that the norms of the objects on the right hand side are $L^p_{\text{loc}}(\Omega')$ functions. The second term is trivial as the point-wise norms $|Df^{-1}(y)|$ are uniformly bounded almost everywhere (remember $f^{-1} \in W^{1,\infty}(\Omega', \mathbb{R}^n)$) and by our hypothesis $|D^3 f| \in L^p_{\text{loc}}(\Omega)$ and the f^{-1}-bilipschitz change of variables does not effect this.

We know by Theorem 2.4 that if $|D^2 f| \in L^q_{\text{loc}}(\Omega)$ then $|D^2 f^{-1}| \in L^q_{\text{loc}}(\Omega')$. The first and the last two terms are essentially the same. We have two bounded factors of $|Df^{-1}|$ and two factors with the same integrability as $|D^2 f|$. We now apply Theorem 2.6 by choosing $i \in \{1, 2, \ldots, n\}$ and taking $u = D_i f$, $r = p$, $q = \infty$, $\hat{k} = 2$, $j = 1$, $\alpha = \frac{1}{2}$ and get $|D^2 f| \in L^{2p}_{\mathrm{loc}}(\Omega)$. Hereby we get that each of the three terms in question are in $L^p_{\mathrm{loc}}(\Omega')$ because clearly $\frac{1}{\frac{1}{2p}+\frac{1}{2p}} = p$. This both proves that we can use the chain and product rule (see Lemma 2.3) and thus $D^3 f^{-1} \in L^p_{\mathrm{loc}}(\Omega')$, which was our claim.

We now continue to the induction step, where we consider $m \geq 4$ and assume that $|D^k f| \in L^q_{\mathrm{loc}}(\Omega)$ implies $|D^k f^{-1}| \in L^q_{\mathrm{loc}}(\Omega')$ for $1 \leq k \leq m-1$ and any $q \in [1, \infty]$. We take the equation (7) and differentiate it $(m-2)$-times, again initially formally and verifying the hypothesis later. The differences here compared to the preliminaries $D^m(f \circ g)$ derivatives are as follows. Instead of equation (1) we start with equation (8) and we proceed similarly with the help of Theorem 2.2. We have an extra Df^{-1} factor, whose derivatives are summed with the components of the Df factor, and therefore we adjust the set $\mathcal{K}_m$ to the set $\mathcal{K}'_m$ as follows

$$\begin{aligned}\mathcal{K}'_m = \Big\{&(k_0, \ldots, k_{m+1}) \in \{0, 1, \ldots, m\}^{m+2} : \\ &(k_0 \geq 2)\ \&\ (k_i = 0 \Leftrightarrow i > k_0 + 1)\ \&\ \Big(\sum_{i=1}^{m+1} k_i = m+1\Big)\Big\}.\end{aligned}$$

We take the extra Df^{-1} factor to correspond to the index 1. The set of orderings Y_k will also differ slightly from X_k as follows,

$$\begin{aligned} Y_k \ni \chi : \{1, \ldots, m\} &\to \{1, \ldots, k_0 + 1\} \\ \min\{j \leq m : \chi(j) = i+1\} &> \min\{j \leq m : \chi(j) = i\} \quad \text{for all } 2 \leq i \leq k_0 \\ \#\{j : \chi(j) = i\} &= k_i \qquad\qquad\qquad\qquad\qquad \text{for all } 2 \leq i \leq k_0 + 1 \\ \#\{j : \chi(j) = 1\} &= k_1 - 1. \end{aligned}$$

Deriving according to Lemma 2.3 and Lemma 3.3 analogously as we did in (5) and (6) we get by Theorem 2.2

$$D^m f^{-1}(y) = \sum_{k \in \mathcal{K}'_m, \chi \in Y_k} D^{k_0} f(f^{-1}(y)) \prod_{i=1}^{k_0+1} D^{k_i} f^{-1}(y) \quad \text{for almost all } y \in \Omega'. \tag{10}$$

The calculations are almost identical to those leading up to (4) and so we omit the details.

Clearly $k_i \leq m-1$ for all $i \geq 1$ and thereby our induction hypothesis tells us that

$$|D^{k_i} f| \in L^q_{\mathrm{loc}}(\Omega) \Rightarrow |D^{k_i} f^{-1}| \in L^q_{\mathrm{loc}}(\Omega') \tag{11}$$

for all $1 \leq i \leq k_0 + 1$ and any $q \in [1, \infty]$. We take the norms of the derivatives from (10) and estimate their integrability. We have $(k_0 + 1)$ factors, which are in the spaces $L^{p_i}_{\text{loc}}(\Omega'), i = 0, 1, \dots, k_0$. We use (11), boundedness of the first order derivatives and Theorem 2.6 (with $u = D_o f$ for $o \in \{1, 2, \dots, n\}$, $r = p$, $j = k_i - 1$, $\hat{k} = m - 1$ and $\alpha = \frac{k_i - 1}{m-1}$) to get that the product in (10) is in the Lebesgue space $L^q_{\text{loc}}(\Omega')$ where

$$\begin{aligned} \frac{1}{q} &= \sum_{i=0}^{k_0+1} \frac{1}{p_i} \\ \frac{1}{p_i} &= \frac{k_i - 1}{n} + \frac{k_i - 1}{m-1}\Big(\frac{1}{p} - \frac{m-1}{n}\Big) = \frac{k_i - 1}{p(m-1)} \\ \frac{1}{q} &= \frac{k_0 - 1}{p(m-1)} + \frac{\sum_{i=1}^{k_0+1}(k_i - 1)}{p(m-1)} = \frac{k_0 - 1}{p(m-1)} + \frac{m - k_0}{p(m-1)} = \frac{1}{p}. \end{aligned} \tag{12}$$

Hence $q = p$. This implies that the norm of the expression in (10) is in $L^p_{\text{loc}}(\Omega')$ and therefore our use of the chain rule and the products rule is correct (see Lemma 2.3 and Theorem 2.2). More significantly, it also implies that $|D^m f^{-1}| \in L^p_{\text{loc}}(\Omega')$, which was the first part of our claim.

Now let us return to BV-regularity of the inverse and assume further that $D^m f \in \mathrm{BV}_{\text{loc}}\big(\Omega, \mathbb{R}^{n^{m+1}}\big)$. Hence $|D^m f| \in L^1_{\text{loc}}(\Omega)$ and from the previous result we know that $f \in W^{m,1}_{\text{loc}}(\Omega, \mathbb{R}^n)$ implies $|D^m f^{-1}| \in L^1_{\text{loc}}(\Omega)$. Moreover, we have

$$D^m f^{-1}(y) = \sum_{k \in \mathcal{K}'_m, \chi \in Y_k} D^{k_0} f(f^{-1}(y)) \prod_{i=1}^{k_0+1} D^{k_i} f^{-1}(y) \quad \text{for almost all } y \in \Omega' \tag{13}$$

and we need to show that the right hand side is a BV function, i.e. its derivative exists and it is a measure. All the terms with $k_0 < m$ can be dealt with as in the previous part of the proof with the help of the Gagliardo-Nirenberg inequality for BV functions Theorem 2.7 and we obtain that the part of the sum with $k_0 < m$ belongs even to $W^{1,1}_{\text{loc}}\big(\Omega', \mathbb{R}^{n^{m+1}}\big)$. It remains to consider the term

$$D^m f(f^{-1}(y)) \prod_{i=1}^{m+1} D^{k_i} f^{-1}(y) = D^m f(f^{-1}(y)) \big(Df^{-1}(y)\big)^{m+1}.$$

Take any $\tilde{\Omega} \subset\subset \Omega$ and corresponding $f(\tilde{\Omega}) = \tilde{\Omega}' \subset\subset \Omega'$. Let us define

$$a_l(y) := D^m f_l(f^{-1}(y)) \big(Df^{-1}(y)\big)^{m+1},$$

where f_l denotes the convolution approximations of f, i.e. $D^m(f_l) \to D^m(f)$ in $L^1(\tilde{\Omega})$ and hence also $D^m(f_l(f^{-1}(y))) \to D^m(f(f^{-1}(y)))$ in $L^1\big(\tilde{\Omega}', \mathbb{R}^{n^{m+1}}\big)$ as f is bilipschitz. Moreover, $D^{m+1} f_l$ (and hence also $D^{m+1} f_l(f^{-1})$) form a bounded sequence in $L^1\big(\tilde{\Omega}', \mathbb{R}^{n^{m+1}}\big)$ and therefore it is not difficult to show using the chain rule that

$$Da_l \text{ is a bounded sequence in } L^1\big(\tilde{\Omega}', \mathbb{R}^{n^{m+1}}\big).$$

We select a subsequence still denoted as Da_l which converges w^* to some Radon measure μ. For all $\varphi \in \mathcal{D}(\tilde{\Omega}')$ we get by the definition of the weak derivative

$$\begin{aligned}\int_{\tilde{\Omega}'} Da_l(y)\varphi(y)\,dy &= -\int_{\tilde{\Omega}'} a_l(y)D\varphi(y)\,dy\\ &= -\int_{\tilde{\Omega}'} D^m f_l(f^{-1}(y))\left(Df^{-1}(y)\right)^{m+1}\, D\varphi(y)\,dy\,.\end{aligned}$$

The left-hand side converges to $\int \varphi\, d\mu$ and hence in the limit we have

$$\int_{\tilde{\Omega}'} \varphi(y)\,d\mu(y) = -\int_{\tilde{\Omega}'} D^m f(f^{-1}(y))\left(Df^{-1}(y)\right)^{m+1} D\varphi(y)\,dy$$

since $D^m(f_l(f^{-1}(y))) \to D^m(f(f^{-1}(y)))$ in $L^1\left(\tilde{\Omega}', \mathbb{R}^{n^{m+1}}\right)$. Clearly our derivative μ is defined uniquely on any open set $G \subset\subset \Omega'$, as for any two $\tilde{\Omega}_1, \tilde{\Omega}_2 \supset G$ and any $\varphi \in \mathcal{D}(G)$ we have $\int_G \varphi d\mu_1 = \int_G \varphi d\mu_2$. This shows that the remaining term (13) belongs to $\mathrm{BV}_{\mathrm{loc}}\left(\Omega', \mathbb{R}^{n^{m+1}}\right)$ and finishes the proof. □

Remark 3.4. To prove the necessity of our assumptions in Theorem 1.2 let us consider the composition of two mappings $f \circ g$, with $f, g : \mathbb{R}^n \to \mathbb{R}^n$.To see that the inverse of the interior must be a Lipschitz mapping consider g to be the projection of $\mathbb{R}^n$ to $\mathbb{R}e_1$ and some f which is not measurable on $\mathbb{R}e_1$.

The next example shows that if f is not Lipschitz the composition may fail to have the original degree of integrability even for bilipschitz g. Consider $n = 9$, $p = \frac{3}{2}$, $\varepsilon \in (0, \frac{1}{3})$,

$$g(x) = x + x|x|^{1+\varepsilon}\sin(|x|^{-1}) \quad \text{and} \quad f(x) = x|x|^{\varepsilon-3}$$

with $\varepsilon > 0$. Main part of the functions is of the form $\frac{x}{|x|}\varphi(|x|)$ and hence we can compute the derivative in a standard way (see e.g. [8, Lemma 2.1]). We know,

$$|D^j f(x)| \approx \frac{f(x)}{|x|^j} \leq |x|^{\varepsilon-2-j} \tag{14}$$

giving us that $f \in W^{4,p}(\Omega, \mathbb{R}^n)$. Further we may estimate the j-th derivative by

$$|D_{1,\ldots,1}g(x)| \geq C\frac{|x|^{2+\varepsilon}}{|x|^{2j}} \tag{15}$$

for some C independent of x for a set with positive density at the origin (i.e. on $S := \{x :\ |\sin(|x|^{-1})| > \frac{1}{4} \text{ and } |\cos(|x|^{-1})| > \frac{1}{4}\}$). From (14) and (15) we get that $f, g \in W^{4,p}(B(0,1))$. We may calculate that

$$|D^4(f \circ g)(x)| \approx |Df(g(x))| \cdot |D^4 g(x)| \approx \frac{1}{|x|^{9-2\varepsilon}} > \frac{1}{|x|^8}$$

on a set with positive density at the origin. Hereby we see that $f \circ g \in W^{4,1}(B(0,1))$ but $D^4 f \circ g \notin L^p(B(0,1))$ as $8p > 9$.

Remark 3.5. The optimal assumptions for the first order regularity of the inverse $f^{-1} \in W^{1,q}$ usually contain some sort of the assumption about the integrability of the distortion function (see e.g. [4, 8] or [10, Chapter 6]).

One can ask if the bilipschitz assumption can be replaced by some condition on the distortion. It is evident that only very restricted results could hold. For example the radial stretching

$$f(x) = \frac{x}{|x|}|x|^{\alpha}$$

with non-zero $\alpha \in \mathbb{R}$ has bounded distortion, as does its inverse. Chose $n, m \in \mathbb{N}$ and put $\alpha = m + 1$. Although $f \in W^{m,\infty}(\Omega, \mathbb{R}^n)$ it can be calculated that $|D^m f^{-1}|^{\beta} \in L^1_{\mathrm{loc}}(\Omega')$ only for

$$\beta < \frac{n}{m - (m+1)^{-1}}.$$

In the previous example we did not have f^{-1} Lipschitz. Nevertheless the counterexample in [7] is a homeomorphism with f^{-1} Lipschitz and the distortion function satisfies $K \in L^{cn}(\Omega)$ so even given these assumptions we can barely expect any a priori results without f bilipschitz.

4. The algebra of bilipschitz Sobolev mappings and its a to smooth Riemannian manifolds

First let us show that bilipschitz Sobolev mappings form an algebra, i.e. they are closed under composition, multiplication and inverse. We start with the following lemma.

Lemma 4.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $g : \Omega' \to \Omega$ be such that $g \in W^{k,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^n) \cap \mathrm{Bilip}_{\mathrm{loc}}(\Omega', \mathbb{R}^n)$ and let $f \in W^{k,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$. Then*

$$\left|D\big(D^{k-1} f(g)\big)\right| \in L^p_{\mathrm{loc}}(\Omega')$$

and

$$D_{a_k}(D_{a_{k-1},\dots,a_1} f(g(x))) = \sum_{l=1}^{n} D_{l,a_{k-1},\dots,a_1} f(g(x)) \big(D_{a_k} g(x)\big)^l$$

or more simply

$$D(D^{k-1} f(g(x))) = D^k f(g(x)) Dg(x).$$

Proof. The proof is similar to that of Lemma 3.3. Instead of the assumption that f^{-1} is bilipschitz, we use that g is bilipschitz and the rest of the reasoning is the same. $\square$

Proof of Theorem 1.2. For $m = 1$ our result follows directly from Theorem 2.2 and the fact that g is bilipschitz. Now let $m = 2$ and use Lemma 4.1 to get that

$$D^2 f \circ g = D^2 f(g) Dg Dg + Df(g) D^2 g.$$

But $|Df(g)|$ and $|Dg|$ are both bounded almost everywhere and $|D^2 f(g)|$ and $|D^2 g|$ are both in $L^p(\Omega')$, therefore our claim holds.

Assume $|D^j f \circ g| \in L^1_{\text{loc}}(\Omega')$, for $j \leq m - 1$. By Lemma 4.1, Theorem 2.2 and Lemma 2.3 repeatedly, we get

$$D^m f \circ g(y) = \sum_{k \in \mathcal{K}_m, \chi \in X_k} D^{k_0} f(g(y)) \prod_{i=1}^{k_0} D^{k_i} g(y)$$

for almost all $y \in \Omega'$, given that the expression on the right is in $L^1_{\text{loc}}(\Omega')$. This can be shown by the calculations corresponding to those in (12) to get that, $D^m(f \circ g) \in L^p(\Omega')$. □

Proof of Theorem 1.3. Let $1 \leq j \leq m - 1$. We calculate using Theorem 2.6 similarly as we did in (12) that $|D^j f|, |D^j g| \in L^{q_j}_{\text{loc}}(\Omega)$, where

$$\frac{1}{q_j} = \frac{j-1}{p(m-1)}.$$

Therefore using the Hölder inequality we get that the product $|D^j f D^{m-j} g| \in L^q_{\text{loc}}(\Omega)$ where

$$\frac{1}{q} = \frac{j-1}{p(m-1)} + \frac{m-j-1}{p(m-1)} < \frac{j-1}{p(m-1)} + \frac{m-1-(j-1)}{p(m-1)} = \frac{1}{p}.$$

Consider the two remaining cases $gD^m f$ and $fD^m g$. These are both clearly in $L^p_{\text{loc}}(\Omega)$. Therefore we can derive fg m-times by the product rule and get that $D^m(fg) \in L^p_{\text{loc}}(\Omega)$. □

It is possible to show that similar result holds also on $\mathcal{C}^\infty$ compact n-dimensional Riemannian manifolds.

Theorem 4.2. *Let M and N be $\mathcal{C}^\infty$ compact n-dimensional connected Riemannian manifolds. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and*

$$\begin{aligned} f &\in W^{m,p}(M, N) \cap \operatorname{Bilip}_{\text{loc}}(M, N), \\ u &\in W^{m,p}(M, N) \cap \operatorname{Lip}(M, N), \\ \varphi &\in W^{m,p}(M, M) \cap \operatorname{Bilip}_{\text{loc}}(M, M) \quad \textit{and} \\ g, h &\in W^{m,p}(M, \mathbb{R}) \cap \operatorname{Lip}(M, \mathbb{R}). \end{aligned}$$

Then

$$\begin{aligned} f^{-1} &\in W^{m,p}(N, M), \\ u \circ \varphi &\in W^{m,p}(M, N) \quad \textit{and} \\ gh &\in W^{m,p}(M, \mathbb{R}). \end{aligned}$$

It is necessary we clarify the meaning of $W^{k,p}(M,N)$ (see e.g. [6]). To begin with we explain that $W^{1,\infty}(M,N) = \mathrm{Lip}(M,N)$ and $W^{1,\infty}(M,\mathbb{R}) = \mathrm{Lip}(M,\mathbb{R})$. Let ρ be the induced metric of the compact Riemannian manifold M. Firstly, notice that all $u \in \mathrm{Lip}(M,\mathbb{R})$ satisfy the following Poincaré type inequality on every ball $B(z,r) \subset M$ and for every $x \in B$,

$$|u(x) - u_B| \leq \fint_B |u(x) - u(y)| \leq \mathrm{Lip}_u \fint_B \rho(x,y) \leq cr\, \mathrm{Lip}_u\,.$$

Notice however that conversely for any mapping u satisfying $|u(x) - u_B| \leq cr$ where the constant c depends on u but not on $B(z,r) \subset M$, is in the class $\mathrm{Lip}(M,\mathbb{R})$. Take any z such that $x, y \in B(z, 3\rho(x,y))$ (given $\rho(x,y) < \infty$) and calculate

$$|u(x) - u(y)| \leq |u(x) - u_B| + |u(y) - u_B| \leq c\rho(x,y).$$

Further it holds, if we have $f : M \to N$, ϕ a map on N and χ a map on M, that

$$\phi \circ f \circ \chi^{-1} \in W^{1,\infty}(U,\mathbb{R}^n) \Leftrightarrow f \in \mathrm{Lip}(M,N),$$

where U is the open set where χ^{-1} is defined. This is because maps are in fact bilipschitz mappings between the manifold and $\mathbb{R}^n$.

Since we have no problems with continuity or in the spaces $W^{1,p}$(M, N) we may use the classical definition below. We will assume without loss of generality that for our $\mathcal{C}^\infty$ compact n-dimensional Riemannian manifolds we have the finite reference atlases $\{\chi_1, \chi_2, \ldots, \chi_k\}, \chi_i : M \to \mathbb{R}^n$, and $\{\phi_1, \phi_2, \ldots, \phi_k\}, \phi_i : N \to \mathbb{R}^n$. Using a division of unity we can conclude that $f \in W^{k,p}(M,N) \cap W^{1,\infty}(M,N)$ if and only if

$$\phi_j \circ f \circ \chi_i^{-1} \in W^{k,p}(U_i,\mathbb{R}^n) \cap W^{1,\infty}(U_i,\mathbb{R}^n)$$

for all $1 \leq i, j \leq k$, where U_i is the open set where χ_i^{-1} is defined.

Proof of Theorem 4.2. Without loss of generality $f(M) = N$. Taking our reference atlases $\{\chi_1, \chi_2, \ldots, \chi_k\}$ on M, and $\{\phi_1, \phi_2, \ldots, \phi_k\}$ on N and by using Theorem 1.1 we obtain that

$$\chi_i \circ f^{-1} \circ \phi_j^{-1} \in W^{k,p}_{\mathrm{loc}}(V_j,\mathbb{R}^n) \cap W^{1,\infty}_{\mathrm{loc}}(V_j,\mathbb{R}^n),$$

where V_j is the open set where ϕ_i^{-1} is defined. Since our reference atlases are finite we easily get that $f^{-1} \in W^{k,p}(N,M) \cap W^{1,\infty}(N,M)$. Following a similar argument as above and calculating

$$\phi_j \circ u \circ \ \varphi \circ \chi_i^{-1}(x) = \phi_j \circ u \circ \chi_l^{-1} \circ \chi_l \circ \varphi \circ \chi_i^{-1}(x)$$

for such x that the expression on the right is defined. Both $\phi_j \circ u \circ \chi_l^{-1}$ and $\chi_l \circ \varphi \circ \chi_i^{-1}$ are $W^{k,p}$-maps where defined. Therefore their composition is also a $W^{k,p}_{\mathrm{loc}}$-map where defined according to Theorem 1.2. The compactness of M again means that the integrability of $u \circ \varphi$ is global. Apply a similar argument to g and h using Theorem 1.3 to get the last result. $\square$

5. The implicit function theorem

In this section we prove a theorem analogous to the **implicit mapping theorem**. Before stating it, let us define some sets. Let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ be open and $u : \Omega \to \mathbb{R}^n$. Then we define

$$\Omega_z^x = \{x \in \mathbb{R}^d : \exists y \in \mathbb{R}^n : (x, y) \in \Omega, u(x, y) = z\} \quad \text{and}$$
$$\Omega^y = \{y \in \mathbb{R}^n : \exists x \in \mathbb{R}^d : (x, y) \in \Omega\}.$$

We will consider mappings which are bilipschitz “in the second variable”. We include our definition here.

Definition 5.1. Let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^m$. The space $\mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ is the class of mappings $u : \Omega \to \mathbb{R}^m$ such that for all $(x_0, y_0) \in \Omega$, $x_0 \in \mathbb{R}^d$, $y_0 \in \mathbb{R}^n$, there exists some $\delta > 0$ and $C_1, C_2 > 0$ such that for all $(x, y), (x, y') \in \Omega \cap B((x_0, y_0), \delta)$ it holds that

$$C_1|y - y'| < |u(x, y) - u(x, y')| < C_2|y - y'|.$$

Theorem 5.2. *Let $k, n, d \in \mathbb{N}, p \in [1, \infty]$, let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ be open and*

$$u \in W^{k,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Then for all $z \in u(\Omega)$, Ω_z^x is open in $\mathbb{R}^d$ and for all $x \in \mathbb{R}^d, y, z \in \mathbb{R}^n$ such that $u(x, y) = z$ there exists a neighbourhood $U_x \subset \Omega_z^x$ of x and $V_y \subset \Omega_y$ of y and exactly one mapping $f_z : U_x \to V_y$ such that

$$u(x', y') = z \Leftrightarrow f_z(x') = y' \quad \textit{for all } (x', y') \in U_x, \times V_y$$

Further given such a triplet x, y, z we have that $f_z \in W^{1,\infty}(U_x, \mathbb{R}^n)$ and for almost all $z \in u(U_x, V_y)$ we have

$$f_z \in W^{k,p}(U_x, \mathbb{R}^n).$$

Remark 5.3. We cannot expect that our hypothesis will guarantee $W^{k,p}$ regularity for every value of z. This can be seen by considering the following function,

$$u(x, y) = y + (x^2 + y^2)^\alpha \sin\left(\frac{1}{x^2 + y^2}\right)$$

with $\alpha \in (2, \frac{5}{2})$ and $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ extended continuously at the origin. Since $u(x, y) - y$ is a radial mapping it is easy to calculate that the norm of the derivative for $j \geq 2$ is

$$|D^j u(x, y)| \leq C \frac{|(x, y)|^{2\alpha}}{|(x, y)|^{3j}}$$

for $|(x,y)| > 0$ and that reverse inequality with different C holds on a a set with positive density at the origin. Notice that u fulfills the hypothesis of Theorem 5.2 with $k=2$ and any $p<\frac{1}{3-\alpha}$. The derivative of u has a singularity only at the origin, which lies on the graph of the corresponding implicit function f_0.

We want to prove that $\int_0^1 |f_0''(s)| = \infty$ and thus $f_0 \notin W^{2,1}$. Set

$$s_m = \frac{1}{\sqrt{\pi m}}$$

and note that $u(s_m, 0) = 0$ which implies that $f_0(s_m) = 0$. We have that

$$\int_0^1 |f_0''(s)| = c + \lim_{j\to\infty} \sum_{m=1}^{j} \int_{s_{m+1}}^{s_m} |f_0''(s)| \geq \lim_{j\to\infty} \sum_{m=1}^{j} |f_0'(s_m+1) - f_0'(s_m)|.$$

The classical implicit function theorem gives that $f|_{(0,1)} \in \mathcal{C}^\infty(0,1)$ and allows us to calculate

$$f_0'(s_m) = -\frac{D_1 u(s_m, 0)}{D_2 u(s_m, 0)} = (-1)^m 2 s_m^{2\alpha-4+1},$$

where we used $\sin\left(\frac{1}{s_m^2+0^2}\right) = 1$ and $\cos\left(\frac{1}{s_m^2+0^2}\right) = (-1)^m$. Now $\alpha < \frac{5}{2}$ implies

$$\lim_{j\to\infty} \sum_{m=1}^{j} |f_0'(s_m+1) - f_0'(s_m)| \geq C \lim_{j\to\infty} \sum_{m=1}^{j} \left(\frac{1}{m\pi}\right)^{\alpha-\frac{3}{2}} + \left(\frac{1}{(m+1)\pi}\right)^{\alpha-\frac{3}{2}} = \infty$$

and therefore $f \notin W^{2,1}((0,1))$. In fact we have $f \notin W^{2,1}((-\delta,\delta))$ for any $\delta > 0$.

We start by proving the following lemma.

Lemma 5.4. *Let* $k, n, d \in \mathbb{N}, p \in [1,\infty]$ *and*

$$u \in W^{k,p}_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n) \cap \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n)$$

Further let $r_1, r_2, L_1 > 0$, $x_0 \in \mathbb{R}^d, y_0 \in \mathbb{R}^n$ *be such that*

$$|u(x,y) - u(x,y')| > L_1 |y - y'| \quad \textit{for all } y, y' \in B_{\mathbb{R}^n}(y_0, r_2) \tag{16}$$

and any $x \in B_{\mathbb{R}^d}(x_0, r_1)$. *Let* $z \in u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{L_1 r_2}{2L_2}\Big)\Big)$ *then there exists* $\delta > 0$ *and a mapping* $f_z : B_{\mathbb{R}^d}(x_0,\delta) \to B_{\mathbb{R}^n}(y_0, r_2)$ *such that*

$$u(x,y) = z \Leftrightarrow f_z(x) = y \quad \textit{for all } x \in B_{\mathbb{R}^d}(x_0, \delta).$$

Further $f_z \in W^{1,\infty}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^n)$ *for all* $z \in u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{L_1 r_2}{2L_2}\Big)\Big)$ *and*

$$f_z \in W^{k,p}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^n) \quad \textit{for almost all } z \in u\left(x_0, B_{\mathbb{R}^n}\left(y_0, \frac{L_1 r_2}{2L_2}\right)\right).$$

Proof. We have $u \in \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n)$ and therefore there exists some $L_2 > 0$ such that

$$|u(x,y) - u(x',y')| < L_2(|x-x'| + |y-y'|) \tag{17}$$

for all $x, x' \in B_{\mathbb{R}^n}(x_0, r_1)$, and all $y, y' \in B_{\mathbb{R}^n}(y_0, r_2)$. Put

$$\delta = \frac{r_2 L_1}{2L_2}.$$

We define $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d \times \mathbb{R}^n$ as follows

$$h(x,y) = (x, u(x,y)).$$

Evidently $h \in W^{k,p}_{\mathrm{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^{d+n})$. We want to prove that

$$h \in \mathrm{Bilip}(B_{\mathbb{R}^d}(x_0, r_1) \times B_{\mathbb{R}^n}(y_0, r_2), \mathbb{R}^{d+n}).$$

It is evident that h is Lipschitz as its component mappings are Lipschitz. Consider $x, x' \in B_{\mathbb{R}^n}(x_0, r_1)$, and $y, y' \in B_{\mathbb{R}^n}(y_0, r_2)$, such that

$$\frac{|y-y'|}{|x-x'|} \leq \frac{2L_2}{L_1}.$$

We have

$$|h(x,y) - h(x',y')| \geq |x-x'| \geq c(|x-x'| + |y-y'|)$$

for some $c > 0$. Now conversely take

$$\frac{|y-y'|}{|x-x'|} \geq \frac{2L_2}{L_1}$$

and by (16) and (17) we get

$$|h(x,y) - h(x',y')| \geq L_1|y-y'| - L_2|x-x'| \geq \frac{L_1|y-y'|}{2} \geq c(|x-x'| + |y-y'|).$$

for some $c > 0$. Now we may denote $\Omega' = h(B_{\mathbb{R}^d}(x_0, r_1), B_{\mathbb{R}^n}(y_0, r_2))$. Clearly Ω' is open. By Theorem 1.1 we get the regularity

$$h^{-1} \in W^{k,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^{d+n}) \cap \mathrm{Bilip}(\Omega', \mathbb{R}^{d+n}).$$

Our goal is to define $f_z(x) = y$. Let us firstly show that if such a y exists then it is unique. The inequality (16) guarantees that if $u(x,y) = u(x,y')$ then $y = y'$. Therefore if $h(x,y) = h(x,y')$ then $y = y'$, which implies that for any given $x \in B_{\mathbb{R}^d}(x_0, r_1)$ and $z \in u(x, B_{\mathbb{R}^n}(y_0, r_2))$ there exists at most one y such that $u(x,y) = z$.

It now suffices to prove that for all $z \in \mathbb{R}^n$ such that $z = u(x_0, \hat{y})$ for some $\hat{y} \in B_{\mathbb{R}^n}\left(y_0, \frac{r_2 L_1}{2L_2}\right)$ we have: for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$ there exists a $y \in B_{\mathbb{R}^n}(y_0, r_2)$

such that $z = u(x, y)$. Remember that $\delta L_2 = \frac{r_2 L_1}{2}$ and put $z_0 = u(x_0, y_0)$. Using the fact that u is L_2-Lipschitz in the y variable and then the definition of δ we get $u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{r_2 L_1}{2L_2}\Big)\Big) \subset B_{\mathbb{R}^n}\Big(z_0, \frac{L_1 r_2}{2}\Big) = B_{\mathbb{R}^n}(z_0, L_1 r_2 - \delta L_2)$. Now since u is L_2-Lipschitz in the x variable and using (16), we get that for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$ that

$$B_{\mathbb{R}^n}(z_0, L_1 r_2 - \delta L_2) \subset B_{\mathbb{R}^n}((x, y_0), L_1 r_2) \subset u\big(x, B_{\mathbb{R}^n}(y_0, r_2)\big).$$

Hence we can really find $y \in B_{\mathbb{R}^n}(y_0, r_2)$ such that $u(x, y) = u(x_0, y_0)$. This gives us the existence of a mapping f_z defined for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$. We show that f_z is Lipschitz (with the constant $\frac{L_2}{L_1}$). Consider two pairs (x, y) and (x', y') such that $u(x, y) = u(x', y') = z$. By (16) and (17) we have

$$\begin{aligned} |u(x, y) - u(x', y)| &< L_2|x - x'| \\ |u(x, y) - u(x', y)| = |u(x', y') - u(x', y)| &> L_1|y - y'|. \end{aligned}$$

Thus

$$|f_z(x) - f_z(x')| = |y - y'| \leq \frac{L_2|x - x'|}{L_1}.$$

It is now left to prove that for almost all $z \in u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{r_2 L_1}{2L_2}\Big)\Big)$ we have $f_z \in W^{k,p}_{\mathrm{loc}}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^n)$. Here it suffices to use Theorem 2.8 and the ensuing comment on the mapping h^{-1} and realize that the ACL condition implies that

$$h^{-1}(\cdot, z) \in W^{k,p}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^{d+n}) \quad \text{for almost all } z \in u\left(x_0, B_{\mathbb{R}^n}\left(y_0, \frac{r_2 L_1}{2L_2}\right)\right).$$

Take any such a point z and use the following notation for the coordinate mappings in the given dimensions d and n, $h^{-1} = (h_1^{-1}, h_2^{-1})$, then clearly $h_2^{-1}(\cdot, z) \in W^{k,p}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^n)$. But clearly for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$ it holds that

$$h_2^{-1}(x, z) = f_z(x).$$

Thus we have $f_z \in W^{k,p}_{\mathrm{loc}}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^n)$ for almost all $z \in u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{r_2 L_1}{2L_2}\Big)\Big)$. □

Proof of Theorem 5.2. We have $u \in \mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ and therefore for any fixed $(x, y) \in \Omega$ we find $r_1, r_2 > 0$ for which we may apply Lemma 5.4 (note that our proof does not require u defined outside of $B(x_0, r_1) \times B(y_0, r_2)$). This means that for any fixed $z \in u(\Omega)$ that Ω^x_z is open in $\mathbb{R}^d$. It also implies the local existence of a Lipschitz $f_z : U_x \to V_y$ for all x, y, z and that for almost all z we have $f_z \in W^{k,p}(U_x, \mathbb{R}^n)$. □

Acknowledgement. The authors would like to thank Robert Černý for carefully reading the manuscript and for his pointed comments. The first two authors were supported by the ERC CZ grant LL1203 of the Czech Ministry of Education.

References

[1] Ambrosio, L., Fusco, N. and Pallara, D., *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford Math. Monogr. New York: Oxford Univ. Press 2000.

[2] Arnold, V., Sur la géometrie differentielle des groupes de Lie de dimension infinie et ses applications á l'hydrodynamique des fluids parfaits (in French). *Ann. Inst. Fourier* 16 (1966), 319 – 361.

[3] Brezis, H. and Mironescu, P., Gagliardo-Nirenberg, composition and products in fractional Sobolev spaces. *J. Evol. Equ.* 1 (2001), 387 – 404.

[4] Csörnyei, M., Hencl, S. and Malý, J., Homeomorphisms in the Sobolev space $W^{1,n-1}$. *J. Reine Angew. Math.* 644 (2010), 221 – 235.

[5] Hajlasz, P., Iwaniec, T., Malý, J. and Onninen, J., Weakly differentiable mappings between manifolds. *Mem. Amer. Math. Soc.* 192 (2008), no. 899.

[6] Hajlasz, P., Sobolev mappings between manifolds and metric spaces. In: *Sobolev Spaces in Mathematics. I.* (ed.: V. Mazya). Int. Math. Ser. (N.Y.) 8. New York: Springer 2009, pp. 185 – 222.

[7] Hencl, S., Bilipschitz mappings with derivative of bounded variation. *Publ. Mat.* 52 (2008), 91 – 99.

[8] Hencl, S. and Koskela, P., *Lectures on Mappings of Finite Distortion.* Lect. Notes Math. 2096. Cham: Springer 2014.

[9] Inci, H., Kappeler, T. and Topalov, P., On the regularity of the composition of diffeomorphisms. *Mem. Amer. Math. Soc.* 226 (2013), no. 1062.

[10] Iwaniec, T. and Martin, G., *Geometric Function Theory and Non-Linear Analysis.* Oxford Math. Monogr. New York: Oxford Univ. Press 2001.

[11] Lee, J. M., *Manifolds and Differential Geometry.* Grad. Stud. Math. 107. Providence (RI): Amer. Math. Soc. 2009.

[12] Mazya, V. and Shaposhnikova, T., An elementary proof of the Brezis and Mironescu theorem on the composition operator in fractional Sobolev spaces. *J. Evol. Equ.* 2 (2002), 113 – 125.

[13] Nirenberg, L., On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa* 3 (1959), 115 – 162.

Received June 25, 2014; revised April 17, 2015