WEAK REGULARITY OF THE INVERSE UNDER MINIMAL ASSUMPTIONS

STANISLAV HENCL, AAPO KAURANEN, AND RAMI LUISTO

Abstract. Let $\Omega \subset \mathbb{R}^3$ be a domain and let $f \in BV_{\rm loc}(\Omega, \mathbb{R}^3)$ be a homeomorphism such that its distributional adjoint is a finite Radon measure. We show that its inverse has bounded variation $f^{-1} \in BV_{\rm loc}$. The condition that the distributional adjoint is finite measure is not only sufficient but also necessary for the weak regularity of the inverse.

1. Introduction

Suppose that $\Omega \subset \mathbb{R}^n$ is an open set and let $f\colon \Omega \to f(\Omega) \subset \mathbb{R}^n$ be a homeomorphism. In this paper we address the issue of the weak regularity of f^{-1} under regularity assumptions on f.

The classical inverse function theorem states that the inverse of a C^1-smooth homeomorphism f is again a C^1-smooth homeomorphism, under the assumption that the Jacobian J_f is strictly positive. In this paper we address the question whether the inverse of a Sobolev or BV-homeomorphism is a BV function or even a Sobolev function. This problem is of particular importance as Sobolev and BV spaces are commonly used as initial spaces for existence problems in PDE's and the calculus of variations. For instance, elasticity is a typical field where both invertibility problems and Sobolev (or BV) regularity issues are relevant (see e.g. [2], [4] and [27]).

The problem of the weak regularity of the inverse has attracted a big attention in the past decade. It started with the result of [18] and [21] where it was shown that for homeomorphisms in dimension $n = 2$ we have

$$\big(f \in W^{1,1}_{\rm loc}(\Omega, \mathbb{R}^2) \text{ is a mapping of finite distortion } \Rightarrow f^{-1} \in W^{1,1}_{\rm loc}(f(\Omega), \mathbb{R}^2)\big)$$
$$\text{and } \big(f \in BV_{\rm loc}(\Omega, \mathbb{R}^2) \Rightarrow f^{-1} \in BV_{\rm loc}(f(\Omega), \mathbb{R}^2)\big).$$

This result has been generalized to $\mathbb{R}^n$ in [6] where it was shown

$$\big(f \in W^{1,n-1}_{\rm loc}(\Omega, \mathbb{R}^n) \text{ is a mapping of finite distortion } \Rightarrow f^{-1} \in W^{1,1}_{\rm loc}(f(\Omega), \mathbb{R}^n)\big)$$
$$\text{and } \big(f \in W^{n-1}_{\rm loc}(\Omega, \mathbb{R}^n) \Rightarrow f^{-1} \in BV_{\rm loc}(f(\Omega), \mathbb{R}^n)\big).$$

It is natural to study the sharpness of the assumption $f \in W^{1,n-1}$. It was shown in [20] that the assumption $f \in W^{1,n-1-\varepsilon}$ (or any weaker Orlicz-Sobolev assumption) is

2010 *Mathematics Subject Classification.* 26B10, 30C65, 46E35.

Key words and phrases. Bounded variation, distributional Jacobian, regularity.

SH and AK were supported in part by the ERC CZ grant LL1203 of the Czech Ministry of Education. SH was supported in part by the grant GAČR P201/18-07996S. AK acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the "María de Maeztu" Programme for Units of Excellence in R&D (MDM-2014- 0445). The last author was supported by a grant of the Finnish Academy of Science and Letters.

not sufficient in general. Furthermore, by results of [8] we know that for $f \in W^{1,n-1}$ we have not only $f^{-1} \in BV$ but also the total variation of the inverse satisfies

$$|Df^{-1}|(f(\Omega)) = \int_{\Omega} |\operatorname{adj} Df(x)|\, dx \tag{1.1}$$

where $\operatorname{adj} Df$ denotes the adjoint matrix, i.e. the matrix of $(n-1)\times(n-1)$ subdeterminants. However, for $n \geq 3$ it is possible to construct a $W^{1,1}$ homeomorphism with $\operatorname{adj} Df \in L^1$ such that $f^{-1} \notin BV$ (see [20]) so the pointwise adjoint does not carry enough information about the regularity of the inverse. The main trouble in the example from [20] is that the pointwise adjoint does not capture some singular behavior on the set of measure zero.

On the other hand, the results of [6] are not perfect as they cannot be applied to even very simple mappings. Let $c(x)$ denote the usual Cantor ternary function, then $h(x) = x + c(x)$ is BV homeomorphism and its inverse $g = h^{-1}$ is even Lipschitz. It is easy to check that

$$f(x,y,z) = [h(x), y, z]$$

is a BV homeomorphism and its inverse $f^{-1}(x,y,z) = [g(x), y, z]$ is Lipschitz, but the results of [6] cannot be applied as f is not Sobolev. In this paper we obtain a new result in dimension $n = 3$ about the regularity of the inverse which generalizes the result of [6] and can be applied to the above mapping.

It is well-known that in models of Nonlinear Elasticity and in Geometric Function Theory the usual pointwise Jacobian does not carry enough information about the mapping and it is necessary to work with the distributional Jacobian, see e.g. [2], [5], [19], [22], [23] and [25]. This distributional Jacobian captures the behavior on zero measure sets and can be used to model for example cavitations of the mapping, see e.g. [15], [16], [26] and [27]. In the same spirit we introduce the notion of the distributional adjoint $\mathcal{ADJ}\, Df$ (see Definition 1.4 below) and we show that the right assumption for the regularity of the inverse is that $\mathcal{ADJ}\, Df \in \mathcal{M}(\Omega, \mathbb{R}^{3\times 3})$, where $\mathcal{M}(\Omega)$ denotes finite Radon measures on Ω. Our main result is the following.

Theorem 1.1. *Let $\Omega \subset \mathbb{R}^3$ be a domain and $f \in BV_{\mathrm{loc}}(\Omega, \mathbb{R}^3)$ be a homeomorphism such that $\mathcal{ADJ}\, Df \in \mathcal{M}(\Omega, \mathbb{R}^{3\times 3})$. Then $f^{-1} \in BV_{\mathrm{loc}}(f(\Omega), \mathbb{R}^3)$.*

If we moreover know that the image of the measure $f(\mathcal{ADJ}\, Df)$ is absolutely continuous with respect to Lebesgue measure, then $f^{-1} \in W^{1,1}_{\mathrm{loc}}(f(\Omega), \mathbb{R}^3)$.

It is not difficult to see that for homeomorphisms in $W^{1,n-1}$ the distributional adjoint $\mathcal{ADJ}\, Df$ is equal to the pointwise adjoint $\operatorname{adj} Df$ (see [19, Proposition 2.10]) and therefore our result generalizes the result of [6] in dimension $n = 3$. The main part of the proof in [6] was to show that f maps $\mathcal{H}^{n-1}$ null sets on almost every hyperplane to $\mathcal{H}^{n-1}$ null sets. This key property may fail in our setting so our proof is more subtle and we have to use delicate tools of Geometric Measure Theory.

Moreover, the condition $\mathcal{ADJ}\, Df \in \mathcal{M}$ is not only sufficient but also necessary for the weak regularity of the inverse.

Theorem 1.2. *Let $\Omega \subset \mathbb{R}^3$ be a domain and $f \in BV(\Omega, \mathbb{R}^3)$ be a homeomorphism such that $f^{-1} \in BV(f(\Omega), \mathbb{R}^3)$. Then $\mathcal{ADJ}\, Df \in \mathcal{M}(\Omega, \mathbb{R}^{3\times 3})$.*

Now we give the formal definition of the distributional adjoint. Without loss of generality we can assume that $\Omega = (0,1)^3$ as all statements are local.

Definition 1.3. *Let $f = (f_1, f_2, f_3)\colon (0,1)^3 \to \mathbb{R}^3$ be a homeomorphism in BV. For $t \in (0,1)$ we define*

$$f_1^t(x) = f(t, x_2, x_3),\ \ f_2^t(x) = f(x_1, t, x_3) \text{ and } f_3^t(x) = f(x_1, x_2, t).$$

We can split these mappings into 9 mappings from $(0,1)^2 \to \mathbb{R}^2$ using its coordinate functions. Given $k, j \in \{1,2,3\}$ choose $a, b \in \{1,2,3\} \setminus \{j\}$ with $a < b$ and define

$$f_{k,j}^t(x) = \big[(f_k^t)_a(x), (f_k^t)_b(x)\big],$$

see Figure 1. For example,

$$f_{1,1}^t(x_2, x_3) = \big[(f_1^t)_2(x), (f_1^t)_3(x)\big] = \big[f_2(t, x_2, x_3), f_3(t, x_2, x_3)\big].$$

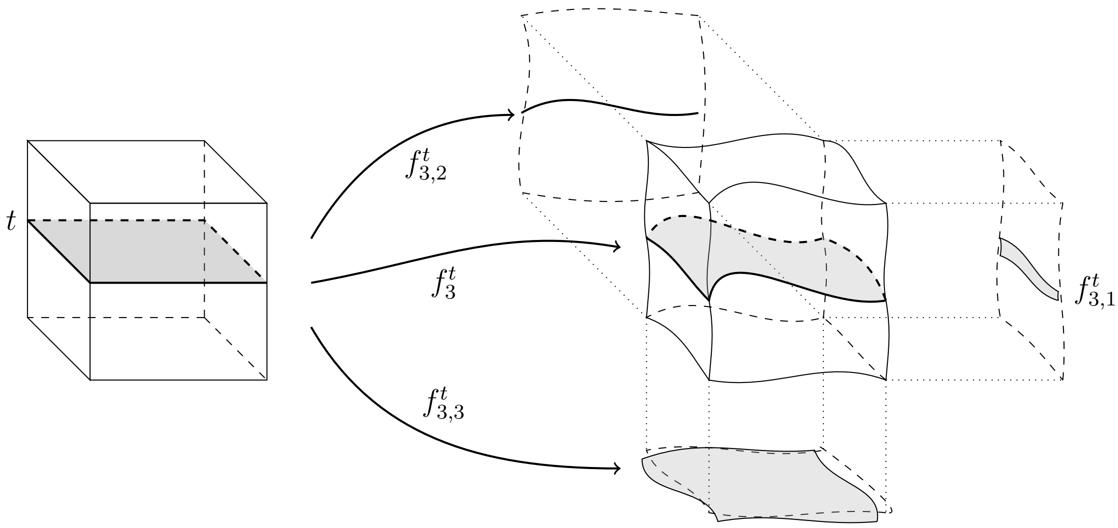

FIGURE 1. The definition of the nine 2-dimensional restrictions $f_{k,j}^t$ of a mapping $f\colon (0,1)^3 \to \mathbb{R}^3$ in Definition 1.3.

Now we recall the definition of distributional Jacobian and, using it, define the distributional adjoint.

Definition 1.4. *Let f be as in Definition 1.3. For mappings $f_{k,j}^t$ we consider the usual distributional Jacobian (see e.g.* [19, Section 2.2]*), i.e. the distribution*

$$\mathcal{J}_{f_{k,j}^t}(\varphi) = -\int_\Omega (f_{k,j}^t)_1(x) J\big(\varphi, (f_{k,j}^t)_2\big)(x)\, dx \qquad \text{for all } \varphi \in C_0^\infty(\Omega).$$

This distribution is well-defined for homeomorphism in $W^{1,1}$. It is also well-defined for homeomorphism in BV for $n = 3$, we just consider the integral with respect to corresponding measure $d(\partial_l f_{k,j}^t)_2(x)$ instead of $(\partial_l f_{k,j}^t)_2(x)\, dx$ and for example we define

$$\begin{aligned}\mathcal{J}_{f_{1,1}^t}(\varphi) = &-\int_\Omega f_2(t, x_2, x_3)\frac{\partial \varphi(x_2, x_3)}{\partial x_2} d\Big(\frac{\partial f_3(t, x_2, x_3)}{\partial x_3}\Big)\\ &+\int_\Omega f_2(t, x_2, x_3)\frac{\partial \varphi(x_2, x_3)}{\partial x_3} d\Big(\frac{\partial f_3(t, x_2, x_3)}{\partial x_2}\Big).\end{aligned}$$

Assume that these 3×3 *distributions* $\mathcal{J}_{f^t_{i,j}}$ *are measures for a.e.* $t\in(0,1)$ *and for measurable* $A\subset(0,1)^3$ *we set*

$$(\mathcal{ADJ}\, Df)_{k,j}(A) = \int_0^1 \mathcal{J}_{f^t_{k,j}}\left(A\cap\{x_k=t\}\right)\, dt.$$

We say that $\mathcal{ADJ}\, Df\in\mathcal{M}(\Omega,\mathbb{R}^{3\times 3})$ *if the distributions* $\mathcal{J}_{f^t_{k,j}}$ *are measures for a.e.* $t\in(0,1)$ *and* $(\mathcal{ADJ}\, Df)_{k,j}\in\mathcal{M}(\Omega)$ *for every* $i,j\in\{1,2,3\}$*.*

A priori it seems that the definition of the distributional adjoint is dependent on the choice of coordinates. This turns out not to be the case and we discuss this further in Section 6.1.

2. Preliminaries

Total variation of the measure μ is the measure $|\mu|$ such that

$$|\mu|(A) := \sup\left\{\int_{\mathbb{R}^n}\varphi\, d\mu:\ \varphi\in C_0(A),\ \|\varphi\|_\infty\le 1\right\} \text{ for all open sets } A\subset\mathbb{R}^n.$$

For a domain $\Omega\subset\mathbb{R}^n$ we denote by $C_0^\infty(\Omega)$ those smooth functions φ whose support is compactly contained in Ω, i.e. $\operatorname{supp}\varphi\subset\subset\Omega$.

2.1. **Mollification.** We will need to approximate continuous BV mappings with smooth maps. To this end we recall here the basic definitions of convolution and mollifiers for the reader's convenience; for a more detailed treatise on the basics and connections to BV mappings we refer to [1, Sections 2.1 and 3.1].

A family of mappings $(\rho_\varepsilon)\in C^\infty(\mathbb{R}^n,\mathbb{R})$ is called *a family of mollifiers* if for all $\varepsilon>0$ we have $\rho_\varepsilon(x)=\varepsilon^{-n}\rho(x/\varepsilon)$, where $\rho\in C^\infty(\mathbb{R}^n,\mathbb{R})$ is a non-negative mapping satisfying $\operatorname{supp}\rho\subset B(0,1)$, $\rho(-x)=\rho(x)$ and $\int_{\mathbb{R}^n}\rho=1$. We will sometimes use a *sequence* of mollifiers (ρ_j), in which case we tacitly assume that there is a family of mollifiers $(\tilde\rho_\varepsilon)$ from which we extract the sequence (ρ_j) by setting $\rho_j:=\tilde\rho_{\frac{1}{j}}$.

For $\Omega\subset\mathbb{R}^n$ and any two functions $f\colon\Omega\to\mathbb{R}^m$, $g\colon\Omega\to\mathbb{R}$ we set their *convolution* to be

$$(f*g)\colon\mathbb{R}^n\to\mathbb{R}^m,\quad (f*g)(x)=\int_\Omega f(y)g(x-y)\, dy,$$

whenever the integral exists. Likewise for a m-valued Radon measure μ defined on Ω and a function $g\colon\Omega\to\mathbb{R}$ we define their convolution as

$$(\mu*f)\colon\mathbb{R}^n\to\mathbb{R}^m,\quad (\mu*f)(x)=\int_\Omega f(x-y)\, d\mu(y)$$

whenever the integral exists.

For a function $f\colon\Omega\to\mathbb{R}^m$ or a Radon measure μ defined on Ω we define their *(family of) mollifications* to be the families $(f_\varepsilon):=(f*\rho_\varepsilon)$ and $(\mu_\varepsilon):=(\mu*\rho_\varepsilon)$, respectively, where (ρ_ε) is a family of mollifiers. Similarly we define the sequence of mollifications as $(f_j):=(f*\rho_j)$ and $(\mu_j):=(\mu*\rho_j)$. For our purposes the exact family of mollifiers does not matter, so we tacitly assume that some such family has been given whenever we use mollifications.

2.2. **Topological degree.** For $\Omega \subset \mathbb{R}^n$ and a given smooth map $f\colon \Omega \to \mathbb{R}^n$ we define the *topological degree* as

$$\deg(f, \Omega, y_0) = \sum_{x \in \Omega \cap f^{-1}\{y_0\}} \operatorname{sgn}(J_f(x))$$

if $J_f(x) \neq 0$ for each $x \in f^{-1}(y_0)$. This definition can be extended to arbitrary continuous mappings and each point $y_0 \notin f(\partial\Omega)$, see e.g. [12, Section 1.2] or [19, Chapter 3.2]. For our purposes the following property of the topological degree is crucial; see [12, Definition 1.18].

Lemma 2.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $f\colon \Omega \to \mathbb{R}^n$ a continuous function and U a domain with $\overline{U} \subset \Omega$. Then for any point $y_0 \in \mathbb{R}^n \setminus f\partial U$ and any continuous mapping $g\colon \Omega \to \mathbb{R}^n$ with*

$$\|f - g\|_\infty \leq \operatorname{dist}(y_0, f(\partial U)),$$

we have $\deg(f, \Omega, y_0) = \deg(g, \Omega, y_0)$.

We will also need some classical results concerning the dependence of the degree on the domain. The following result is from [12, Theorem 2.7].

Lemma 2.2. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $f\colon \Omega \to \mathbb{R}^n$ a continuous function and U a domain with $\overline{U} \subset \Omega$.*

(1) (Domain decomposition property) *For any domain $D \subset U$ with a decomposition $D = \cup_i D_i$ into open disjoint sets, and a point $p \notin f\partial D$, we have*
$$\deg(f, D, p) = \sum_i \deg(f, D_i, p).$$
(2) (Excision property) *For a compact set $K \subset \overline{U}$ and a point $p \notin f(K \cup \partial U)$ we have* $\deg(f, U, p) = \deg(f, U \setminus K, p)$.

The topological degree agrees with the Brouwer degree for continuous mappings, which in turn equals the *winding number* in the plane. The winding number is an integer expressing how many times the path $\beta_f := f\partial D$ circles the point p; indeed, the winding number equals the topological index of the mapping $\frac{\beta_f - p}{|\beta_f - p|}\colon \mathbb{S}^1 \to \mathbb{S}^1$. We refer to [12, Section 2.5] for discussion of the winding number in the setting of holomorphic planar mappings.

2.3. **Hausdorff measure.** For $A \subset \mathbb{R}^n$ we use the classical definition of the Hausdorff measure (see e.g. [10])

$$\mathcal{H}^k(A) = \lim_{\delta \to 0+} \mathcal{H}^k_\delta(A),$$

where

$$\mathcal{H}^k_\delta(A) = \inf\Big\{\sum_i \operatorname{diam}^k A_i : \ A \subset \bigcup_i A_i, \ \operatorname{diam} A_i \leq \delta\Big\}.$$

The important ingredient of our proof is the Gustin boxing inequality [13] which states that for each compact set $K \subset \mathbb{R}^n$ we have

$$\mathcal{H}^{n-1}_\infty(K) \leq C_n \mathcal{H}^{n-1}(\partial K). \tag{2.1}$$

2.4. **On various areas.** Besides homeomorphisms in three dimensions we work with a continuous mappings $g\colon [0,1]^2 \to \mathbb{R}^3$ in our paper. Besides the usual Hausdorff measure of the image $\mathcal{H}^2(g((0,1)^2))$ we need to use other finer notions of area. The results of this subsection can be found in the book of Cesari [3] and they also follow by some results of Federer, see e.g. [11, (13) on page 93] and references given there.

First we define the Lebesgue area (see [3, 3.1]). Let L be an affine mapping, then for any triangle Δ the area of $L(\Delta)$ is defined in the natural way. For a piecewise linear mapping $h : [0,1]^2 \to \mathbb{R}^3$ we define the Lebesgue area $A(h,[0,1]^2)$ to be the sum of areas of triangles of some triangulation where h is linear in each of these triangles. We define

$$A(g,[0,1]^2) := \inf\left\{\lim_{k\to\infty} A(g_k([0,1]^2)) : (g_k) \in \mathcal{PH}(g)\right\}, \tag{2.2}$$

where $\mathcal{PH}(g)$ is the collection of all sequences of polyhedral surfaces converging uniformly to g.

Next we define a coordinate mappings $g_j\colon [0,1]^2 \to \mathbb{R}^2$ as

$$g_1(x) = [g_2(x), g_3(x)],\ \ g_2(x) = [g_1(x), g_3(x)] \text{ and } g_3(x) = [g_1(x), g_2(x)].$$

Finally we define (see [3, 9.1])

$$V(g_j,[0,1]^2) := \sup_S \left\{ \sum_{\pi\in S} \int_{\mathbb{R}^2} \left|\deg(g_j,\pi,y)\right|\, dy \right\}, \tag{2.3}$$

where S is any finite system of nonoverlapping simple open polygonal regions in $[0,1]^2$ and $\deg(g_j,y,A)$ denotes the topological degree of mapping.

We need the following characterization of the Lebesgue area (see [3, 18.10 and 12.8.(ii)]) which holds for any continuous g

$$V(g_j,[0,1]^2) \le A(g,[0,1]^2) \le V(g_1,[0,1]^2) + V(g_2,[0,1]^2) + V(g_3,[0,1]^2). \tag{2.4}$$

Let us note that these results are highly nontrivial. For example it is possible to construct continuous g such that $A(g,[0,1]^2)$ is much smaller than $\mathcal{H}^2(g([0,1]^2))$ (which may be even infinite) but the result (2.4) is still true. Further for the validity we need only continuity of g and we do not need to assume that $A(g,[0,1]^2) < \infty$. However, this is only known to hold for two dimensional surfaces in $\mathbb{R}^3$ and for higher dimensions the assumption about the finiteness of the Lebesgue area might be needed.

2.5. BV **functions and the coarea formula.** Let $\Omega \subset \mathbb{R}^n$ be an open domain. A function $h \in L^1(\Omega)$ is of bounded variation, $h \in BV(\Omega)$, if the distributional partial derivatives of h are measures with finite total variation in Ω, i.e. there are Radon (signed) measures $\mu_1,\dots,\mu_n$ defined in Ω so that for $i = 1,\dots,n$, $|\mu_i|(\Omega) < \infty$ and

$$\int_\Omega hD_i\varphi\, dx = -\int_\Omega \varphi\, d\mu_i$$

for all $\varphi \in C_0^\infty(\Omega)$. We say that $f \in L^1(\Omega,\mathbb{R}^n)$ belongs to $BV(\Omega,\mathbb{R}^n)$ if the coordinate functions of f belong to $BV(\Omega)$.

Let $\Omega \subset \mathbb{R}^n$ be open set and $E \subset \Omega$ be a measurable. The perimeter of E in Ω is defined as a total variation of χ_E in Ω, i.e.

$$P(E,\Omega) := \sup\left\{\int_E \operatorname{div}\varphi\, dx :\ \varphi \in C_0^1(\Omega),\ \|\varphi\|_\infty \le 1\right\}.$$

We will need the following coarea formula to characterize BV functions (see [1, Theorem 3.40]).

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in L^1(\Omega)$. Then we have*

$$|Du|(\Omega) = \int_{-\infty}^{\infty} P(\{x \in \Omega : \ u(x) > t\}, \Omega)\, dt. \tag{2.5}$$

In particular, $u \in BV(\Omega)$ if and only if the integral on the righthand side is finite.

It is well-known (see e.g. [1, Proposition 3.62]) that for a coordinate functions of a homeomorphism $f\colon \Omega \to \mathbb{R}^n$ we have

$$P\big(\{x \in \Omega : \ f_i(x) > t\}, \Omega\big) \le \mathcal{H}^{n-1}\big(\{x \in \Omega : \ f_i(x) = t\}\big). \tag{2.6}$$

Moreover, we have the following version of coarea formula for continuous BV functions by Federer [10, Theorem 4.5.9 (13) and (14) for $k \equiv 1$].

Theorem 2.4. *Let $\Omega \subset \mathbb{R}^n$ be open and $u \in BV(\Omega)$ be continuous. Then we have*

$$|Du|(\Omega) = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}(\{x \in \Omega : \ u(x) = t\})\, dt.$$

2.6. **BVL condition.** Let $\Omega \subset \mathbb{R}^n$ be open and $f \in L^1(\Omega)$. It is well-known that $f \in BV(\Omega)$ if and only if it satisfies the BVL condition, i.e. it has bounded variation on $\mathcal{L}^{n-1}$ a.e. line parallel to coordinate axes, and the variation along these lines is integrable (see e.g. [1, Remark 3.104]). As a corollary we obtain that BV function of n-variables is a BV function of $(n-1)$-variables on $\mathcal{L}^1$ a.e. hyperplane parallel to coordinate axis.

For example for $n = 2$ and $f \in BV((0,1)^2)$ we have that the function

$$f_x(y) := f(x, y)$$

has bounded (one-dimensional) variation for a.e. $x \in (0,1)$. Moreover,

$$\int_0^1 |Df_x((0,1))|\, dx = |D_2 f|((0,1)^2) \tag{2.7}$$

where $|Df_x|$ denotes the (one-dimensional) total variation of f_x and $|D_2 f|$ denotes the total variation of the measure $\frac{\partial f}{\partial y}$. Similar identity holds for $f_y(x) := f(x,y)$ and $D_1 f$.

2.7. **Convergence of BV functions.** In dimension two, the boundary of a ball $B(x,r)$ is a curve and we will tacitly assume that it is always parametrized with the path

$$\beta\colon [0, 2\pi] \to \mathbb{R}^2, \quad \beta(t) = (x_1 + r\cos t, x_2 + r \sin t).$$

Thus when we speak of the *length* $\ell(\partial B(x,r))$ of the boundary of a ball or its image $f\partial B(x,r)$ under a mapping f, we mean the length of the curve β or $f \circ \beta$, respectively. Note that the length of a path $\gamma\colon [0,1] \to \mathbb{R}^2$ equals

$$\ell(\gamma) := \left\{ \sum_{j=1}^{k} d(\gamma(t_{j-1}), \gamma(t_j)) : 0 = t_0 \le \ldots \le t_k = 1 \right\},$$

from which we immediately see that if $f_j \to f$ uniformly, then

$$\lim_{j\to\infty} \ell(f_j \circ \beta) \geq \ell(f \circ \beta).$$

Similarly we also assume line segments in the plane to be equipped with a path parametrization and to have similar length convergence properties.

By the results in the previous subsection 2.6 we know that the restriction of a BV mapping $f \in BV(\mathbb{R}^2, \mathbb{R}^2)$ to $\mathcal{L}^1$ a.e. line segment in the plane is again a BV mapping; i.e. for all $a > b$ and a.e. $x \in \mathbb{R}$, the restriction $f_x := f|_{\{x\}\times(a,b)}$ is a BV mapping and

$$\mathcal{H}^1(fI_x) \leq \ell(fI_x) = |Df_x|(I_x) < \infty,$$

where $|Df_x|$ denotes the one-dimensional total variation of f_x.

A similar result holds also for $\mathcal{H}^1$ a.e. radius of a sphere: given a point x, the restriction of f to $\partial B(x,r)$ is BV for $\mathcal{H}^1$ a.e. radius $r > 0$. This especially implies that for such radii,

$$\mathcal{H}^1(f\partial B(x,r)) \leq \ell(f\partial B(x,r)) = |D(f_r)|(\partial B(x,r)) < \infty,$$

where $f_r := f|_{\partial B(x,r)}$ and $|D(f_r)|$ denotes the one-dimensional total variation of f_r. Furthermore, similarly as in (2.7), we have

$$\int_0^r \mathcal{H}^1(f\partial B(x,s))\, ds \leq \int_0^r |D(f_s)|(\partial B(x,s))\, ds \leq |Df|\,(B(x,r))\,. \tag{2.8}$$

Recall the weak* convergence of BV mappings.

Definition 2.5. *We say that a sequence (f_j) of BV mappings* weakly* converges to f in BV, *if $f_j \to f$ in L^1 and Df_j weakly* converge to Df, i.e.*

$$\lim_{j\to\infty} \int_\Omega \varphi\, dDf_j = \int_\Omega \varphi\, dDf$$

for all $\varphi \in C_0(\Omega)$.

The following result from [1, p.125, Proposition 3.13.] gives a characterization for weak* convergence in BV. Note especially that since mollifications of continuous BV functions converge uniformly, they especially converge locally in L^1, so in this case the boundedness of the sequence in BV-norm gives weak* convergence for the derivatives.

Proposition 2.6. *Let f_j be a sequence of BV mappings $\Omega \to \mathbb{R}^2$. Then f_j weakly* converges to a BV mapping $f\colon \Omega \to \mathbb{R}^2$ if and only $f_j \to f$ in L^1 and* $\sup |Df_j|(\Omega) < \infty$.

3. Properties of BV mappings

In the proof of Theorem 1.1 we use some ideas of [8, proof of Theorem 1.7]. In particular we use the following observation based on the coarea formula (Theorem 2.3).

Theorem 3.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain and $f \in BV_{\rm loc}(\Omega, \mathbb{R}^n)$ be a homeomorphism. Then the following measure on Ω is finite*

$$\mu(A) = \sum_{i=1}^n \int_{-\infty}^\infty \mathcal{H}^{n-1}\big(f(\{x \in A:\ x_i = t\})\big)\, dt$$

if and only if $f^{-1} \in BV_{\rm loc}(f(\Omega), \mathbb{R}^3)$*. In addition,* $f(\mu)$ *is absolutely continuous with respect to the Lebesgue measure if and only if* $f^{-1} \in W^{1,1}(f(\Omega), \mathbb{R}^3)$*.*

Proof. Assume that μ is a finite measure. By Theorem 2.3 and the perimeter inequality (2.6) we have

$$
\begin{aligned}
|Df^{-1}|(f(\Omega)) &\approx \sum_{i=1}^{n} |(Df^{-1})_i|(f(\Omega)) \\
&= C \sum_{i=1}^{n} \int_{-\infty}^{\infty} P(\{y \in f(\Omega) : \ (f^{-1})_i(y) > t\}, f(\Omega)) \, dt \\
&\leq C \sum_{i=1}^{n} \int_{-\infty}^{\infty} \mathcal{H}^{n-1}\big(\{y \in f(\Omega) : \ (f^{-1})_i(y) = t\}\big) \, dt \\
&= C \sum_{i=1}^{n} \int_{-\infty}^{\infty} \mathcal{H}^{n-1}\big(f(\{x \in \Omega : \ x_i = t\})\big) \, dt < \infty.
\end{aligned} \tag{3.1}
$$

and thus $f^{-1} \in BV_{\rm loc}$.

If $f^{-1} \in BV_{\rm loc}$, then by Theorem 2.4 we know that the only inequality in the above computation (3.1) is actually equality and we get $\mu \in \mathcal{M}(\Omega)$.

Let us consider now the final claim. We have to show that $|Df^{-1}|(E) < \varepsilon$ if $|E| < \delta$. Given ε we choose $\delta > 0$ from the absolute continuity of measure $f(\mu)$. By approximation we may assume that E is open and $|E| < \delta$. The definition of μ, assumed absolute continuity of $f(\mu)$ and (3.1) (with E instad of Ω) imply

$$\left|Df^{-1}\right|(E) \leq \mu(f^{-1}(E)) < \varepsilon.$$

If we know that $f^{-1} \in W^{1,1}$ then we have only equalities in (3.1) and we easily obtain that $f(\mu)$ is absolutely continuous with respect to the Lebesgue measure. □

We next show that for a mollification of a continuous BV mapping, the convergence is inherited to a.e. circle in a weak sense.

Proposition 3.2. *Let* $f\colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$ *be a continuous* BV *mapping and* (ρ_k) *a sequence of mollifiers. Denote* $f^k := f * \rho_k$*,* $f^k_{c,s} := f^k|_{\partial B(c,s)}$ *and* $f_{c,s} := f|_{\partial B(c,s)}$*. Then for any point* $z \in \mathbb{R}^2$ *we have*

$$\lim_{k\to\infty} |Df^k_{z,r}|(\partial B(z,r)) = |Df_{z,r}|(\partial B(z,r)) < \infty$$

and

$$D(f^k_{z,r}) \overset{w*}{\rightharpoonup} D(f_{z,r}),$$

for $\mathcal{H}^1$ *a.e. radius* $r > 0$ *such that* $\overline{B(z,r)} \subset \Omega$*.*

Proof. Since the claim is local in nature it suffices, after a smooth change of local coordinates, to show that for a continuous BV mapping $f\colon (0,1)^2 \to \mathbb{R}^2$ we have

$$\lim_{k\to\infty} |Df^k_x|(I_x) = |Df_x|(I_x) < \infty, \tag{3.2}$$

and

$$D(f^k_x) \overset{w*}{\rightharpoonup} D(f_x), \tag{3.3}$$

on $\mathcal{H}^1$ almost every line segment $I_x := \{x\} \times (0,1)$, where $f^k_x := f^k|_{I_x}$ and $f_x := f|_{I_x}$.

We start by proving (3.2). By the results in Section 2.7, for $\mathcal{H}^1$ a.e. $x \in (0,1)$ we have

$$|Df_x^k|(I_x) = \ell(f^k I_x) < \infty \quad \text{and} \quad |Df_x|(I_x) = \ell(f I_x) < \infty,$$

so since $f_x^k \to f_x$ uniformly we see by the notions of Section 2.7 that

$$\lim_{k\to\infty} |Df_x^k|(I_x) = \lim_{k\to\infty} \ell(f^k I_x) \geq \ell(f I_x) = |Df_x|(I_x).$$

Thus to prove (3.2) it suffices to show that for a.e. $x \in (0,1)$,

$$\lim_{k\to\infty} |Df_x^k|(I_x) \leq |Df_x|(I_x).$$

Suppose this is not true, whence there exists $\delta > 0$ such that the set

$$J := \Big\{ x \in (0,1) : \lim_{k\to\infty} |Df_x^k|(I_x) > (1+\delta)|Df_x|(I_x) \Big\}$$

has positive 1-measure. Fix a Lebesgue point $x_0 \in (0,1)$ of J. By the Lebesgue density theorem we may assume x_0 to be such that

$$\lim_{r\to 0} \frac{1}{2r} \int_{x_0-r}^{x_0+r} \big| |Df_x|(I_x) - |Df_{x_0}|(I_{x_0}) \big| = 0.$$

Choose $\eta > 0$ such that

$$(1+\delta)\frac{1-\eta}{2}\big((1-\eta)2 - \eta\big) > 1. \tag{3.4}$$

Fix $r > 0$ for which

$$\begin{aligned} &(i)\ |Df|\big(\partial((x_0-r, x_0+r)\times(0,1))\big) = 0, \\ &(ii)\ |J \cap (x_0-r, x_0+r)| \geq (1-\eta)2r, \\ &(iii)\ \int_{x_0-r}^{x_0+r} \big| |Df_x|(I_x) - |Df_{x_0}|(I_{x_0}) \big| \, dx < \eta |Df_{x_0}|(I_{x_0}) r, \text{ and} \\ &(iv)\ |Df_{x_0}|(I_{x_0}) > \frac{1-\eta}{2r} \int_{x_0-r}^{x_0+r} |Df_x|(I_x)\, dx. \end{aligned} \tag{3.5}$$

As remarked in Section 2.6, $D_2 f$ is a finite Radon measure. Thus applying [1, Proposition 2.2.(b), p.42] for the mollification of its total variation $|D_2 f|$ and using the fact that the measure is Borel regular, we see that

$$\lim_{k\to\infty} |D_2 f^k|(U) \leq \lim_{k\to\infty} |D_2 f| \left(U + B(0, k^{-1})\right) = |D_2 f|(\overline{U}).$$

for any Borel set U. By setting $U := (x_0 - r, x_0 + r) \times (0,1)$ we have by (i) that $|Df|(\partial U) = 0$, and so also $|D_2 f|(\partial U) \leq |Df|(\partial U) = 0$. Thus

$$\lim_{k\to\infty} |D_2 f^k|(U) \leq |D_2 f|(\overline{U}) = |D_2 f|(U). \tag{3.6}$$

On the other hand by using Fatou's lemma, definition of J, (3.5) (iii), (ii), (iv) and (3.4),

$$
\begin{aligned}
\lim_{k\to\infty} |D_2 f^k|(U) &\geq \lim_{k\to\infty} \int_{(x_0-r,x_0+r)\cap J} |Df_x^k|(I_x)\, dx \\
&\geq (1+\delta) \int_{(x_0-r,x_0+r)\cap J} |Df_x|(I_x)\, dx \\
&\geq (1+\delta) \Big[\int_{(x_0-r,x_0+r)\cap J} |Df_{x_0}|(I_{x_0})\, dx - \eta |Df_{x_0}|(I_{x_0}) r \Big] \\
&\geq (1+\delta) |Df_{x_0}|(I_{x_0}) \big((1-\eta) 2r - \eta r \big) \\
&\geq (1+\delta) \frac{1-\eta}{2r} \int_{x_0-r}^{x_0+r} |Df_x|(I_x)\, dx \big((1-\eta) 2r - \eta r \big) \\
&> |D_2 f^k|(U).
\end{aligned}
$$

This contradicts (3.6) and so (3.2) holds.

To prove (3.3) we note that for $\mathcal{H}^1$ a.e. line segment I_x the BV mappings $f_x^k\colon I_x \to \mathbb{R}^2$ converge uniformly to the continuous BV mapping $f_x\colon I_x \to \mathbb{R}^2$. Furthermore they form a bounded sequence with respect to the BV norm, and thus by Proposition 2.6 they converge weak* in BV. This implies (3.3) and the proof is complete. □

4. Degree Theorem for continuous BV planar mappings

The aim of this section is to prove the following analogy of the change of variables formula for the distributional Jacobian in two dimensions. Similar statement was shown before in [5] for mappings that satisfy $J_f > 0$ a.e. and that are one-to-one and in [9] for open and discrete mappings. Here we generalize this result to mappings where the Jacobian can change the sign but we restrict our attention to planar mappings only.

Theorem 4.1. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous BV mapping such that the distributional Jacobian $\mathcal{J}f$ is a signed Radon measure. Then for every $x \in \mathbb{R}^2$ we have*

$$
\int_{\mathbb{R}^2} \deg(f, B(x,r), y)\, dy = \mathcal{J}_f\,(B(x,r)) \tag{4.1}
$$

for a.e. $r > 0$.

Before the proof Theorem 4.1 we prove the following important corollary, which is one of the main tools in the proof of Theorem 1.1.

Proposition 4.2. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous BV mapping such that the distributional Jacobian $\mathcal{J}f$ is a signed Radon measure. Then for every $x \in \mathbb{R}^2$ we have*

$$
\int_{\mathbb{R}^2} |\deg(f, B(x,r), y)|\, dy \leq |\mathcal{J}f|\,(B(x,r)) \tag{4.2}
$$

for a.e. $r > 0$.

Proof. Let $x_0 \in \mathbb{R}^2$. By the BVL properties of f we know that for almost every radius $r > 0$ we have $\ell(f\partial B(x,r)) < \infty$. Fix such a radius r and denote $B_0 := B(x,r)$. We enumerate the bounded components of $\mathbb{R}^2 \setminus f\partial B_0$ as $\{U_j\}$, set $W_j :=$

$B(x,r)\cap f^{-1}U_j$ and enumerate the components of W_j as $\{W_j^m\}_m$, and finally denote $C := B_0 \cap f^{-1}f(\partial B_0)$. Note that the degree map $y \mapsto \deg(f, B_0 \setminus C, y)$ is constant in the components U_j; we denote this constant by $\deg(f, B_0 \setminus C, U_j)$. Furthermore, $fC = f\partial B_0$, so especially $\mathcal{L}^2(fC) = 0$, and $\partial W_j^m \subset C$, $f\partial W_j^m \subset \partial U_j$, for all j, m.

Now, on one hand

$$\mathcal{J}_f(B_0) = \mathcal{J}_f(C) + \mathcal{J}_f(B_0 \setminus C) = \mathcal{J}_f(C) + \mathcal{J}_f(\bigcup_{m,j} W_j^m)$$

so especially

$$|\mathcal{J}_f|(B_0) = |\mathcal{J}_f|(C) + \sum_j |\mathcal{J}_f|(W_j). \tag{4.3}$$

On the other hand by using the fact that fC has zero measure, the excision property of the degree (Lemma 2.2 (2)), the domain decomposition property of the degree (Lemma 2.2 (1)) and the fact that $f(W_j^m) \subset U_j$ we have

$$\begin{aligned}
\int_{\mathbb{R}^2} |\deg(f, B_0, y)|\, dy &= \sum_j \int_{U_j} |\deg(f, B_0, y)|\, dy \\
&= \sum_j \int_{U_j} |\deg(f, B_0 \setminus C, y)|\, dy = \sum_j \mathcal{L}^2(U_j)\, |\deg(f, B_0 \setminus C, U_j)| \\
&= \sum_j \left| \mathcal{L}^2(U_j) \sum_m \deg(f, W_j^m, U_j) \right| = \sum_j \left| \int_{U_j} \sum_m \deg(f, W_j^m, y)\, dy \right| \\
&= \sum_j \left| \int_{\mathbb{R}^2} \deg(f, W_j, y)\, dy \right|.
\end{aligned} \tag{4.4}$$

Fix j and denote $U := U_j$ and $W := W_j$. Let us note that the statement of Theorem 4.1 holds not only for a.e. ball but also for a.e. rectangle. Thus we may take a collection of rectangles $\{Q^k\}_k$ such that $\bigcup_k Q^k = W$, $\mathcal{L}^2(\bigcup_k f\partial Q^k) = 0$ and

$$\int_{\mathbb{R}^2} \deg(f, Q^k, y)\, dy = \mathcal{J}_f\left(Q^k\right)$$

for all k. Now similarly as for the calculations of (4.4), we enumerate the bounded components of $\mathbb{R}^2 \setminus f(\cup_k \partial Q^k)$ as $\{\tilde{U}_i\}_i$, again note that the degree map $y \mapsto \deg(f, Q^k, y)$ is constant in the components $\tilde{U}_i$ and denote this constant as $\deg(f, Q^k, \tilde{U}_i)$. With this notation we see that

$$\left| \int_{\tilde{U}_i} \deg(f, Q^k, y)\, dy \right| = \begin{cases} \left|\mathcal{J}_f(Q^k)\right|, & \text{if } f(Q^k) \subset \tilde{U}_i \\ 0, & \text{otherwise,} \end{cases}$$

and the sum over k and i of the left hand side is smaller than $|\mathcal{J}_f|(W)$. Thus the following summation with respect to the indices i and k is interchangeable:

$$\sum_i \sum_k \int_{\tilde{U}_i} \deg(f, Q^k, \tilde{U}_i)\, dy = \sum_k \sum_i \int_{\tilde{U}_i} \deg(f, Q^k, \tilde{U}_i)\, dy$$

and we may calculate that

$$\begin{aligned}\int_{\mathbb{R}^2} \deg(f,W,y)\,dy &= \sum_i \int_{\tilde U_i} \deg(f,W,y)\,dy = \sum_i \int_{\tilde U_i} \sum_k \deg(f,Q^k,y)\,dy\\ &= \sum_i \mathcal{L}^2(\tilde U_i) \sum_k \deg(f,Q^k,\tilde U_i) = \sum_i \sum_k \int_{\tilde U_i} \deg(f,Q^k,y)\,dy\\ &= \sum_k \int_{\mathbb{R}^2} \deg(f,Q^k,y)\,dy = \sum_k \mathcal{J}_f(Q^k).\end{aligned}$$

This in turn then naturally implies that

$$\left|\int_{\mathbb{R}^2} \deg(f,W,y)\,dy\right| \le \sum_k |\mathcal{J}_f(Q^k)| \le |\mathcal{J}_f|(W). \tag{4.5}$$

Thus by combining inequalities (4.3), (4.4) and (4.5) we see that

$$|\mathcal{J}_f|(B_0) \ge |\mathcal{J}_f|(C) + \sum_j |\mathcal{J}_f|(W_j) \ge \sum_j \left|\int_{\mathbb{R}^2} \deg(f,W_j,y)\,dy\right| = \int_{\mathbb{R}^2} |\deg(f,B_0,y)|\,dy$$

and the proof is complete. □

The proof of Theorem 4.1 requires several auxiliary results. We begin with the following degree convergence lemma; compare to Lemma 2.1.

Lemma 4.3. *Let $f\colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous BV mapping and let (f^k) be mollifications of f. Then for any point $x \in \mathbb{R}^2$ and a.e. radius $r > 0$ we have*

$$\lim_{k\to\infty} \int_{\mathbb{R}^2} \deg(f^k, B(x,r), y)\,dy = \int_{\mathbb{R}^2} \deg(f, B(x,r), y)\,dy. \tag{4.6}$$

Proof. Let $x_0 \in \mathbb{R}^2$. By the BVL properties remarked in Section 2.6 we know that for almost every radius $r > 0$ the length of $f\partial B(x_0,r)$ is finite, i.e. $|Df_{z,r}|(\partial B(z,r)) < \infty$ and that the claim of Proposition 3.2 holds. Fix such a r_0, and set $B_0 := B(x_0,r_0)$.

We first define

$$F^k(y) = \deg(f^k, B(x_0,r), y) \quad\text{and}\quad F(y) = \deg(f, B(x_0,r), y),$$

whence

$$\int_{\mathbb{R}^2} \deg(f^k, B(x_0,r_0), y)dy = \left\| F^k \right\|_1 \quad\text{and}\quad \int_{\mathbb{R}^2} \deg(f, B(x_0,r_0), y)dy = \|F\|_1$$

and we have to show that $F^k \to F$ in L^1. To show this we use compactness of BV. First we show that F^k is bounded sequence in BV-norm.

It is easy to see that the variation measure of F^k is supported only on the curve $f^k(\partial B(x_0,r_0))$. Furthermore, since $f^k(\partial B(x_0,r_0))$ is rectifiable, $\mathcal{H}^1$-a.e. point is on the boundary of at most two components of $\mathbb{R}^2 \setminus f^k(\partial B(x_0,r_0))$. In such a situation, if the value of F^k differs by N on these two components, the image $f^k(\partial B(x_0,r_0))$ must cover this joint boundary at least N times. Thus the total variation of F^k is in fact bounded by

$$\ell\big(f^k(\partial B(x_0,r_0))\big) = \left|Df^k_{x_0,r_0}\right|(\partial B(x_0,r_0)).$$

Since the radius r_0 was chosen such that Proposition 3.2 holds, we have

$$\left|Df^k_{x_0,r_0}\right|(\partial B(x_0,r_0)) \to \left|Df_{x_0,r_0}\right|(\partial B(x_0,r_0))$$

and so $\left|DF^k\right|$ is uniformly bounded. Furthermore the boundedness of the sequence (F^k) in L^1 follows from the Sobolev inequality [1, Theorem 3.47]. Thus, the compactness theorem in [1, Theorem 3.23] implies that there exists a subsequence $(F^{k(j)})$ which converges in L^1 to a function G.

We will show that $G = F$, which implies that the original sequence F^k converges to F in L^1, as every converging subsequence must converge to F. Assume that $F \neq G$ on a set A with positive Lebesgue measure. Since $f(\partial B(x_0, r_0))$ has finite 1-Hausdorff measure we find with Lebesgue density theorem $z \in \mathbb{R}^2 \setminus f(\partial B(x_0, r_0))$, which is a density point of A with $G(z) \neq F(z)$. For some very small ball B_z centered at z we have

$$\left| \int_{B_z} G - F \right| > 0$$

and B_z is compactly contained in some component of $\mathbb{R}^2 \setminus f(\partial B(x_0, r_0))$. Now recall that f^k converge uniformly to f. When $\| f^k - f \| < \operatorname{dist}(B_z, f(\partial B(x_0, r_0)))$ we have by basic properties of the degree (see [12, Theorem 2.3.])

$$F^k(y) = \deg(f^k, B, y) = \deg(f, B, y) = F(y)$$

for every point $y \in B_z$. This is a contradiction with L^1 convergence and the definition of B_z. Thus the original claim follows. □

The proof of the previous lemma goes through also with absolute values of the degrees. We record this observation as the following corollary even though we will not be using it in this paper.

Corollary 4.4. *Let $f\colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous BV mapping and (f^k) a mollification of f. Then for any point $x \in \mathbb{R}^2$ and a.e. radius $r > 0$ we have*

$$\lim_{k\to\infty} \int_{\mathbb{R}^2} |\deg(f^k, B(x,r), y)|\, dy = \int_{\mathbb{R}^2} |\deg(f, B(x,r), y)|\, dy. \tag{4.7}$$

The following Proposition 4.5 is essentially a BV-version of [17, Proposition 2.10]. For smooth mappings the identity (4.10) follows in a more general form with smooth test functions $g \in C^\infty(\Omega, \mathbb{R}^2)$ by combining the Gauss-Green theorem and the area formula in a ball B:

$$\begin{aligned} \int_{\partial B} \langle (g(f(x)) \cdot \operatorname{cof} Df(x), \nu \rangle \; d\mathcal{H}^1(x) &= \int_B \operatorname{div} g(f(y)) J_f(y)\, dy \\ &= \int_{\mathbb{R}^2} \operatorname{div} g(y) \deg(f, B, y)\, dy, \end{aligned} \tag{4.8}$$

where ν denotes the unit exterior normal to B and $\operatorname{cof} Df(x)$ denotes the cofactor matrix, i.e. the matrix of $(n-1) \times (n-1)$ subdeterminants with correct signs. For more details for the general setting we refer to Müller, Spector and Tang; in [27, Proposition 2.1] they prove the claim for continuous $f \in W^{1,p}$, $p > n-1$ and $g \in C^1$. We need the identity only in the case of $g(x_1, x_2) = [x_1, 0]$. In this case the integrand on left hand side of (4.8) reduces to

$$f_1 \langle Df_2, \nu_t \rangle\,, \tag{4.9}$$

where ν_t is the unit tangent vector of ∂B. Thus in the BV setting it is natural to replace the left hand side of (4.8) with

$$\int_{\partial B} f_1 d(Df|_{\partial B}).$$

since by Section 2.6 f is one dimensional BV-function on almost every sphere centered at any given point.

Proposition 4.5. *Let $\Omega \subset \mathbb{R}^2$ be a domain and let $f\colon \Omega \to \mathbb{R}^2$ be a continuous BV mapping. Then for every $c \in \mathbb{R}^2$ and a.e. $r > 0$ such that $B := B(c, r) \subset \Omega$ we have*

$$\int_{\partial B} f_1 d(Df|_{\partial B}) = \int_{\mathbb{R}^2} \deg(f, B, y)\, dy. \tag{4.10}$$

Proof. We prove the claim by approximating f with a sequence of mollifiers (f^k), showing that $[f_1^k, 0] \cdot \operatorname{cof} Df^k$ convergences weakly* to $f_1 d(Df|_{\partial B})$ and combining this with Lemma 4.3.

Let us fix $r > 0$ such that $B(c, r) \subset\subset \Omega$ and the conclusion of Lemma 4.3 and Proposition 3.2 hold for this radius. Now for every k with $B(c, r + \frac{1}{k}) \subset\subset \Omega$ we have $f^k \in C^\infty(B, \mathbb{R}^2)$. Since f is continuous, $f^k \to f$ uniformly. Clearly

$$\int_{\partial B} \big(f_1^k Df_2^k|_{\partial B} - f_1 Df|_{\partial B}\big) = \int_{\partial B} \big(f_1 Df_2^k|_{\partial B} - f_1 Df|_{\partial B}\big) - \int_{\partial B} \big(f_1^k - f_1\big) Df_2^k|_{\partial B}.$$

We next note that by Proposition 3.2, $Df^k|_{\partial B} \to Df|_{\partial B}$ with respect to the weak* convergence and $\|f_1^k - f_1\|_\infty \to 0$ by the uniform convergence of (f^k). Thus both terms of the right hand side converge to zero as $k \to \infty$. It follows that

$$\begin{aligned} &\lim_{k\to\infty} \int_{\partial B} \big\langle [f_1^k(x), 0] \cdot \operatorname{cof} Df^k(x), \nu \big\rangle \, d\mathcal{H}^1(x) = \\ &\qquad = \lim_{k\to\infty} \int_{\partial B} f_1^k(x) Df^k|_{\partial B}(x) d\mathcal{H}^1(x) = \int_{\partial B} f_1(x) d(Df|_{\partial B}(x)). \end{aligned} \tag{4.11}$$

On the other hand, since the mappings f^k are smooth we have by e.g. [27, Proposition 2.1] that

$$\int_{\partial B} \big\langle [f_1^k(x), 0] \cdot \operatorname{cof} Df^k(x), \nu \big\rangle \, d\mathcal{H}^1(x) = \int_{\mathbb{R}^2} \deg(f^k, B, y) dy.$$

Combining this with (4.11) and Lemma 4.3 gives the claim. □

We are now ready to prove the main result of this section, Theorem 4.1. In its proof we use some ideas from [5] and [25].

Proof of Theorem 4.1. We recall the definition of distributional Jacobian for any $\varphi \in C_0^\infty(\Omega)$

$$\mathcal{J}_f(\varphi) = -\int_\Omega f_1(x) J(\varphi(x), f_2(x))\, dx = \int_\Omega \big\langle [f_1(x), 0] \cdot \operatorname{cof} Df(x), D\varphi(x) \big\rangle \, dx. \tag{4.12}$$

Let us pick a ball $B := B(y, r) \subset \Omega$ such that $|Df|(\partial B) = 0$. Furthermore, by the Lebesgue theorem we may assume that

$$\lim_{\delta\to 0} \frac{1}{\delta} \int_{r-\delta}^{r} \Big| |Df_{y,s}|(\partial B(y, s)) - |Df_{y,r}|(\partial B(y, r)) \Big| \, ds = 0 \tag{4.13}$$

where $f_{y,s} := f|_{\partial B(y,s)}$ and $|Df_{y,s}|$ is the corresponding (one-dimensional) total variation. Let us fix $\psi \in C^\infty(\mathbb{R},[0,1])$ such that $\psi(s) \equiv 1$ for $s < 0$ and $\psi(s) \equiv 0$ for $s > 1$. For $0 < \delta < r$ we set

$$\Phi_\delta(s) = \psi\Big(\frac{s-(r-\delta)}{\delta}\Big), \text{ i.e. } \Phi_\delta(s) = \begin{cases} 1 \text{ for } s \leq r-\delta. \\ 0 \text{ for } s \geq r. \end{cases} \quad \text{and } |\Phi_\delta'| \leq \frac{C}{\delta}.$$

As the distributional Jacobian is a Radon measure and $|Df|(\partial B) = 0$ we obtain

$$\mathcal{J}_f(B(y,r)) = \lim_{\delta\to 0+} \int_\Omega \Phi_\delta(|x-y|)\, d\mathcal{J}_f(x). \tag{4.14}$$

By (4.12) for $\varphi = \Phi_\delta(|x-y|)$ and Proposition 4.5 we have

$$\begin{aligned} \int_\Omega \Phi_\delta(|x-y|)\, d\mathcal{J}_f(x) &= \int_\Omega \big\langle [f_1(x), 0]\cdot \operatorname{cof} Df(x), D\Phi_\delta(|x-y|)\big\rangle\, dx \\ &= \int_{r-\delta}^r \int_{\partial B(y,s)} f_1(x)\Phi_\delta'(s) d(Df|_{\partial B}(x)) \\ &= \int_{r-\delta}^r \Phi_\delta'(s) \int_{\mathbb{R}^2} \deg(f, B(y,s), z)\, dz\, ds. \end{aligned} \tag{4.15}$$

We next show that the integral on the right hand side of (4.15) converges as $\delta \to 0$. For all $y \in \Omega$ and $\delta > 0$ small enough we set

$$f_s(x) = f\left(\frac{s}{r}(x-y) + y\right).$$

Note that with this notation

$$\int_{r-\delta}^r \Phi_\delta'(s) \int_{\mathbb{R}^2} \deg(f, B(y,s), z)\, dz\, ds = \int_{r-\delta}^r \Phi_\delta'(s) \int_{\mathbb{R}^2} \deg(f_s, B(y,r), z)\, dz\, ds \tag{4.16}$$

and the right hand side is a type of average integral as $\int_{r-\delta}^r \Phi' = 1$.

We have a fixed mapping $f|_{\overline{B(y,r)}}$ with $|Df_{y,r}|(\partial B(y,r)) < \infty$. We claim that given $\varepsilon > 0$ we can find $\eta > 0$ such that for every continuous mapping $g|_{\overline{B(y,r)}}$ we have

$$\begin{aligned} &\|f-g\|_{L^\infty(\partial B)} < \eta \text{ and } \big||Df_{y,r}|(\partial B(y,r)) - |Dg_{y,r}|(\partial B(y,r))\big| < \eta \Rightarrow \\ &\qquad \Rightarrow \Big|\int_{\mathbb{R}^2} \deg(f, B(y,r), z)\, dz - \int_{\mathbb{R}^2} \deg(g, B(y,r), z)\, dz\Big| < \varepsilon. \end{aligned} \tag{4.17}$$

Indeed, if this were not true, we would have uniformly converging sequence such that conclusion of (4.17) would not hold. Analogously to the proof of Lemma 4.3 we would then get a contradiction.

Moreover, similarly to the proof of Lemma 4.3, the Sobolev inequality gives for these a.e. radii

$$\Big|\int_{\mathbb{R}^2} \deg(f_s, B(y,r), z)\, dz\Big| \leq C|Df_{y,s}|(\partial B(y,s)). \tag{4.18}$$

Given $\varepsilon > 0$ we choose $\eta > 0$ as in (4.17) and then we choose $\delta > 0$ so that for every $s \in [r-\delta, r]$ we have

$$\|f-f_s\|_{L^\infty(\partial B)} < \eta \text{ and } \frac{1}{\delta}\int_{r-\delta}^r \Big||Df_{y,s}|(\partial B(y,s)) - |Df_{y,r}|(\partial B(y,r))\Big|\, ds < \eta^2 \tag{4.19}$$

where we have used (4.13). By the Chebyshev's inequality with (4.19) we obtain

$$|W| < \eta\delta \text{ for } W := \Big\{ s \in [r-\delta, r] : \ \big| |Df_{y,s}|(\partial B(y,s)) - |Df_{y,r}|(\partial B(y,r)) \big| > \eta \Big\}.$$

By (4.15), (4.16), $\int_{r-\delta}^{r} \Phi'_\delta = 1$, $|\Phi'_\delta| \le \frac{C}{\delta}$, (4.17) and (4.18) we obtain

$$\begin{aligned}
&\Big| \int_\Omega \Phi_\delta(|x-y|)\, d\mathcal{J}_f(x) - \int_{\mathbb{R}^2} \deg(f, B(y,r), z)\, dz \Big| = \\
&\quad = \Big| \int_{r-\delta}^{r} \Phi'_\delta(s) \Big(\int_{\mathbb{R}^2} \deg(f_s, B(y,r), z)\, dz - \int_{\mathbb{R}^2} \deg(f, B(y,r), z)\, dz \Big)\, ds \Big| \\
&\quad \le \frac{C}{\delta} \Big[\int_{[r-\delta,r]\setminus W} \varepsilon + \int_W \big(|Df_{y,s}|(\partial B(y,s)) + |Df_{y,r}|(\partial B(y,r)) \big)\, ds \Big] \\
&\quad \le C\varepsilon + \frac{C}{\delta} \int_W \big| |Df_{y,s}|(\partial B(y,s)) - |Df_{y,r}|(\partial B(y,r)) \big|\, ds + \frac{2C}{\delta} \int_W |Df_{y,r}|(\partial B(y,r)) ds \\
&\quad \le C\varepsilon + C\eta^2 + 2C\eta |Df_{y,r}|(\partial B(y,r)).
\end{aligned}$$

Together with (4.14) this implies that

$$\mathcal{J}_f(B(y,r)) = \lim_{\delta \to 0+} \int_\Omega \Phi_\delta(|x-y|)\, d\mathcal{J}_f(x) = \int_{\mathbb{R}^2} \deg(f, B(y,r), z)\, dz.$$

□

5. Proof of main Theorem 1.1

Proof of Theorem 1.1. Without loss of generality we may assume that $(-1,2)^3 \subset \Omega$ and we prove only that $f^{-1} \in BV\big(f((0,1)^3)\big)$ as the statement is local. Let us denote $Q := (0,1)^2$ and square with same center and double sidelength is denoted by $2Q$. We claim that

$$\int_0^1 \mathcal{H}^2\big(f(Q \times \{t\})\big) dt < \infty$$

and the statement of the theorem then follows from Theorem 3.1. Let $\varepsilon > 0$. We start with an estimate for $\mathcal{H}^2_\varepsilon(f(Q \times \{t\}))$ for some fixed $t \in (0,1)$.

First let us fix $t \in (0,1)$ such that

$$\lim_{\delta \to 0} \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} \sum_{j=1}^{3} \big| |\mathcal{J}_{f^s_{3,j}}|(2Q \times \{s\}) - |\mathcal{J}_{f^t_{3,j}}|(2Q \times \{t\}) \big|\, ds = 0 \tag{5.1}$$

and we note that this holds for a.e. $t \in (0,1)$ by the Lebesgue density theorem. Let us define the measure on $(0,1)$ by

$$\begin{aligned}
\mu((a,b)) = &\sum_{j=1}^{3} \int_{-1}^{2} |\mathcal{J}_{f^s_{2,j}}|\big((-1,2) \times \{s\} \times (a,b)\big)\, ds \\
&+ \sum_{j=1}^{3} \int_{-1}^{2} |\mathcal{J}_{f^s_{1,j}}|\big(\{s\} \times (-1,2) \times (a,b)\big)\, ds.
\end{aligned} \tag{5.2}$$

Let us denote by h the absolutely continuous part of μ with respect to $\mathcal{L}^1$. Then it is easy to see that

$$\int_0^1 h \leq \mu((0,1)) \leq |\mathcal{ADJ}\, Df|\big((-1,2)^2 \times (0,1)\big). \tag{5.3}$$

Moreover, we can fix t so that

$$\lim_{\delta\to 0} \frac{\mu\big((t-\delta,t+\delta)\big)}{2\delta} = h(t)$$

which holds for a.e. t by the Lebesgue density theorem and by the fact that the corresponding limit is zero a.e. for the singular part of μ.

Since f is uniformly continuous there exists for our fixed t a subdivision of $Q\times\{t\} = \bigcup_i Q_i$ into rectangles such that $\operatorname{diam}(f(2Q_i \times \{t\})) < \frac{\varepsilon}{2}$. Furthermore we fix $r > 0$ so that for every $0 < \delta < r$ we have with the help of (5.1)

$$\begin{aligned}
&(i)\ \frac{\mu\big((t-\delta,t+\delta)\big)}{2\delta} \leq 2h(t),\\
&(ii)\ \sum_{j=1}^3 |\mathcal{J}_{f_{3,j}^{t+\delta}}|(2Q\times\{t+\delta\}) \leq 2\sum_{j=1}^3 |\mathcal{J}_{f_{3,j}^{t}}|(2Q\times\{t\}),\\
&(iii)\ \sum_{j=1}^3 |\mathcal{J}_{f_{3,j}^{t-\delta}}|(2Q\times\{t-\delta\}) \leq 2\sum_{j=1}^3 |\mathcal{J}_{f_{3,j}^{t}}|(2Q\times\{t\}) \text{ and}\\
&(iv)\ \operatorname{diam}\big(f(2Q_i\times[t-\delta,t+\delta])\big) < \varepsilon \text{ for each } i.
\end{aligned} \tag{5.4}$$

For $\eta > 0$ we put our $Q_i \times \{t\}$ into the box

$$U_{i,\eta} := (1+\eta)Q_i \times [t-\delta,t+\delta].$$

In the following we divide $\partial U_{i,\eta}$ into three parts parallel to coordinate axes

$$\partial_3 U_{i,\eta} := (1+\eta)Q_i \times \{t-\delta,t+\delta\},\ \partial_2 U_{i,\eta} \text{ and } \partial_1 U_{i,\eta},$$

where $\partial_2 U_{i,\eta}$ denotes two rectangles perpendicular to x_2 axis and $\partial_1 U_{i,\eta}$ denotes two rectangles perpendicular to x_1 axis. For each Q_i we choose a real number $0 \leq \eta_i \leq 1$ so that

$$\sum_{k=1}^2\sum_{j=1}^3 |\mathcal{J}_{f_{k,j}}|(\partial_k U_{i,\eta_i}) \leq \int_0^1 \sum_{k=1}^2\sum_{j=1}^3 |\mathcal{J}_{f_{k,j}}|(\partial_k U_{i,\eta'})\, d\eta', \tag{5.5}$$

which is possible as the smallest value is less or equal to the average and here and in the following we denote for simplicity $|\mathcal{J}_{f_{1,j}}|(\partial_1 U_{i,\eta})$ the sum of two

$$|\mathcal{J}_{f_{1,j}^{c_1\pm(1+\eta)r}}|\big(\{c_1 \pm (1+\eta)r\}\times[c_2-(1+\eta)r, c_2+(1+\eta)r]\times[t-\delta,t+\delta]\big),$$

where $Q_i = Q(c,r)$.

It is obvious that $f(Q_i\times\{t\}) \subset f(U_{i,\eta_i})$. By the definition of the Lebesgue area (2.2) and its estimate (2.4) we obtain that we can approximate f on $\partial U_{i,\eta_i}$ by piecewise linear $f^i\colon \partial U_{i,\eta_i} \to \mathbb{R}^3$ such that

$$\mathcal{H}^2(f^i(\partial U_{i,\eta_i})) \leq 2A(f,\partial U_{i,\eta_i}) = 2\sum_{k=1}^3 A(f_k,\partial_k U_{i,\eta_i}) \leq 2\sum_{k=1}^3\sum_{j=1}^3 V(f_{k,j},\partial_k U_{i,\eta_i}) \tag{5.6}$$

and so that f^i is so close to f that (see (5.4) (iv))

$$\operatorname{diam}\big(f^i(\partial U_{i,\eta_i})\big) < \varepsilon \tag{5.7}$$

and $f(Q_i \times \{t\})$ lies inside $f^i(U_{i,\eta_i})$, i.e.

$$f(Q_i \times \{t\}) \subset G_i := \bigcup \text{ bounded components of } \mathbb{R}^3 \setminus f^i(\partial U_{i,\eta_i}).$$

By (5.7) we have

$$\mathcal{H}^2_\varepsilon(G_i) = \mathcal{H}^2_\infty(G_i).$$

Now we obtain for each t with the help of Gustin boxing lemma (2.1) and (5.6)

$$\begin{aligned}\mathcal{H}^2_\varepsilon(f(Q\times\{t\})) &\leq \sum_i \mathcal{H}^2_\varepsilon(f(Q_i\times\{t\})) \leq \sum_i \mathcal{H}^2_\varepsilon(G_i)\\ &= \sum_i \mathcal{H}^2_\infty(G_i) \leq C\sum_i \mathcal{H}^2(\partial G_i) \leq C\sum_i \mathcal{H}^2(f^i(\partial U_{i,\eta_i}))\\ &\leq C\sum_i\sum_{k=1}^3\sum_{j=1}^3 V(f_{k,j},\partial_k U_{i,\eta_i}).\end{aligned} \tag{5.8}$$

Recall that by definition (2.3)

$$V(h,U) = \sup_S \left\{ \sum_{\pi\in S} \int_{\mathbb{R}^2} \big|\deg(h,\pi,y)\big|\, dy \right\},$$

so by Proposition 4.2 we obtain

$$\begin{aligned}\sum_{k=1}^2\sum_{j=1}^3 V(f_{k,j},\partial_k U_{i,\eta_i}) &\leq C\sum_{k=1}^2\sum_{j=1}^3 |\mathcal{J}_{f_{k,j}}|(\partial_k U_{i,\eta_i}) \text{ and}\\ \sum_{j=1}^3 V(f_{3,j},\partial_3 U_{i,\eta_i}) &\leq C\sum_{j=1}^3 |\mathcal{J}_{f_{3,j}}|(\partial_3 U_{i,\eta_i}).\end{aligned} \tag{5.9}$$

Notice that even though Theorem 4.2 is stated only for disks, it also holds for rectangles and moreover, we may use it for polygons. This can be seen by covering the polygon by rectangles and arguing as in the end of the proof of Proposition 4.2.

We treat the terms in (5.9) separately. We sum the last inequality, use the fact that $(1+\eta_i)Q_i$ have bounded overlap (as $1 \leq 1+\eta_i \leq 2$) and with the help of (5.4) (ii) and (iii) we obtain

$$\begin{aligned}\sum_i\sum_{j=1}^3 V(f_{3,j},\partial_3 U_{i,\eta_i}) &\leq C\sum_i\sum_{j=1}^3 |\mathcal{J}_{f_{3,j}}|(\partial_3 U_{i,\eta_i})\\ &\leq C\sum_{j=1}^3\Big(|\mathcal{J}_{f_{3,j}}|(2Q\times\{t-\delta\}) + |\mathcal{J}_{f_{3,j}}|(2Q\times\{t+\delta\}\Big)\\ &\leq C\sum_{j=1}^3 |\mathcal{J}_{f_{3,j}}|(2Q\times\{t\}).\end{aligned} \tag{5.10}$$

For the remaining part of the right hand side of (5.8) we denote $Q_i = Q(c,r)$ and by (5.9), (5.5), linear change of variables and $\delta < r$ we have

$$\begin{aligned}\sum_{k=1}^{2}\sum_{j=1}^{3} V(f_{k,j},\partial_k U_{i,\eta_i}) &\leq C\sum_{k=1}^{2}\sum_{j=1}^{3} |\mathcal{J}_{f_{k,j}}|(\partial_k U_{i,\eta_i})\\ &\leq C\int_0^1 \sum_{k=1}^{2}\sum_{j=1}^{3} |\mathcal{J}_{f_{k,j}}|(\partial_k U_{i,\eta'})\, d\eta'\\ &\leq \frac{C}{\delta}\int_{c_1-2r}^{c_1+2r} \sum_{j=1}^{3} |\mathcal{J}_{f^a_{1,j}}|\big(\{a\}\times[c_2-2r,c_2+2r]\times[t-\delta,t+\delta]\big)\, da\\ &+\frac{C}{\delta}\int_{c_2-2r}^{c_2+2r} \sum_{j=1}^{3} |\mathcal{J}_{f^a_{2,j}}|\big([c_1-2r,c_1+2r]\times\{a\}\times[t-\delta,t+\delta]\big)\, da.\end{aligned}$$

Summing over i, using bounded overlap of $2Q_i$, (5.2) and (5.4) (i) we obtain

$$\sum_i\sum_{k=1}^{2}\sum_{j=1}^{3} V(f_{k,j},\partial_k U_{i,\eta_i}) \leq C\frac{\mu\big((t-\delta,t+\delta)\big)}{\delta} \leq Ch(t). \tag{5.11}$$

Combining (5.8), (5.10) and (5.11), we have with the help of (5.3)

$$\begin{aligned}\int_0^1 \mathcal{H}^2_\varepsilon\big(f(Q\times\{t\})\big)\, dt &\leq C\int_0^1 \sum_{j=1}^{3} |\mathcal{J}_{f_{3,j}}|(2Q\times\{t\})\, dt + \int_0^1 h(t)\, dt\\ &\leq C|\mathcal{ADJ}\, Df|((-1,2)^3).\end{aligned}$$

By passing $\varepsilon\to 0$ we obtain our conclusion with the help of Theorem 3.1. □

6. Reverse implication

The main aim of this Section is to show Theorem 1.2. For its proof we again use some ideas from Müller [25] and De Lellis [7]. As a corollary we show that the notion of $\mathcal{ADJ}\, Df \in \mathcal{M}$ does not depend on the chosen system of coordinates and that this notion is weakly closed.

For the proof of Theorem 1.2 we require the following result which shows that the topological degree is smaller than the number of preimages.

Lemma 6.1. *Let $F\colon \mathbb{R}^3\to\mathbb{R}^3$ be a homeomorphism, $f\colon\mathbb{R}^2\to\mathbb{R}^3$ the restriction of F to the xy-hyperplane, $p\colon\mathbb{R}^3\to\mathbb{R}^2$ the projection $(x_1,x_2,x_3)\mapsto(x_1,x_2)$ and $g=p\circ f$. Then for any $B(x,r)\subset(0,1)^2$ and $y\in\mathbb{R}^2\setminus g\partial B(z,r)$,*

$$|\deg(g,B(z,r),y)| \leq N(g,B(z,r),y).$$

Proof. We may assume that $N(g,B(z,r),y)$ is finite and $\deg(g,B(z,r),y)>0$. By [12, Theorem 2.9] we can write the degree as a sum of local indices

$$\deg(g,B(z,r),y) = \sum_{x\in B(z,r)\cap g^{-1}\{y\}} i(g,x,y);$$

recall that local index is defined by

$$i(g,x,y) := \deg(g,V,y),$$

where V is any neighborhood of x such that $g^{-1}(y) \cup \bar{V} = \{x\}$.

To prove the claim it thus suffices to prove that $|i(g,x,y)| \leq 1$ for every $x \in g^{-1}\{y\}$. Towards contradiction suppose that this is not the case. Fix some $x_0 \in g^{-1}\{y\}$ such that $i(g,x,y) \geq 2$; the case when the index is negative is dealt identically. Let $B(x_0,s)$ be a ball such that

$$i(g,x_0,y) = \deg(g, B(x_0,s), y).$$

Without loss of generality we may assume that $x_0 = y = 0$, $s = 1$. Denote $Z = \{0\} \times \{0\} \times \mathbb{R}$. Since the topological degree equals the winding number, the path $\beta := g(\partial B(0,1))$ winds around the point 0 at least twice in $\mathbb{R}^2 \setminus \{0\}$, so especially the path $\alpha := f(\partial B(0,1))$ winds twice around Z in $\mathbb{R}^3 \setminus Z$.

Now we note that $\partial B(0,1) \times \{0\} \subset \mathbb{S}^2$, where $\mathbb{S}^2$ denotes the two-dimensional sphere in $\mathbb{R}^3$. Since $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ is a a homeomorphism and f the restriction of F, $F(\mathbb{S}^2)$ is a topological sphere in $\mathbb{R}^3$. Furthermore, Z intersects $fB(0,1)$ only at a single point, and we fix $\hat{Z}$ to be the compact subinterval of Z which contains the intersection point and intersects $F(\mathbb{S}^2)$ only at the endpoints of the interval, which we may assume to be $(0,0,\pm 1)$. The unique pre-images of these points cannot be on the circle $\partial B(0,1)$, so we may assume then to be $(0,0,\pm 1)$ as well. Thus

$$\alpha = f(\partial B(0,1)) = F(\partial B(0,1) \times \{0\}) \subset F(\mathbb{S}^2) \setminus \hat{Z}.$$

This gives rise to a contradiction, since F is a homeomorphism and so the degree of $F|_{\mathbb{S}^2 \setminus Z}$ is ± 1. More specifically, the path $\alpha\colon \mathbb{S}^1 \to F(\mathbb{S}^2) \setminus Z$ winds around the Z-axis at least twice, i.e. the homotopy class $[\alpha]$ of α in the group $\pi_1(\mathbb{R}^2 \setminus Z, \alpha(0)) \simeq \mathbb{Z}$ is non-zero and does not span the group $\mathbb{Z}$. Furthermore the intersection $Z \cap F(B^3(0,1))$ consists of countably many paths starting and ending at the boundary $f\mathbb{S}^2$ and so since Z intersects $fB(0,1)$ only at a single point all but one of these loops can be pulled to the boundary $f\mathbb{S}^2$ without intersecting α. Thus the homotopy class $[\alpha]$ of α in the group $\pi_1(F(\mathbb{S}^2) \setminus \hat{Z}, \alpha(0)) \simeq \mathbb{Z}$ is also non-zero and does not span the group $\mathbb{Z}$. But this is a contradiction since $\alpha = F(\partial B(0,1) \times \{0\})$, where the homotopy class $[\partial B(0,1) \times \{0\}]$ spans $\pi_1(F(\mathbb{S}^2) \setminus Z, (1,0,0)) \simeq \mathbb{Z}$ at the domain side and a homeomorphism F induces an isomorphism between homotopy groups by e.g. [14, p. 34]. ☐

Proof of Theorem 1.2. The distributional adjoint is a well-defined distribution as $f \in BV$ is continuous. Without loss of generality we assume that f is defined on $(0,1)^3$ and we show that $\mathcal{ADJ}\, Df \in \mathcal{M}((0,1)^3)$. We only show that $\mathcal{J}_{f^t_{1,1}}$ is a measure for a.e. $t \in (0,1)$ and that

$$\int_0^1 \mathcal{J}_{f^t_{1,1}}((0,1)^2)\, dt < \infty \tag{6.1}$$

as the proof for other eight components of $\mathcal{ADJ}\, Df$ is similar.

By Theorem 3.1 we know that

$$\int_0^1 \mathcal{H}^2\big(f^t_1((0,1)^2)\big)\, dt < \infty.$$

Let us fix $t \in (0,1)$ such that $\mathcal{H}^2\big(f^t_1((0,1)^2)\big) < \infty$. Put $g := f^t_{1,1}$ and denote by g_1 and g_2 its coordinate functions. Let us fix $\varphi \in C^1_0((0,1)^2)$. We recall the definition of distributional Jacobian

$$\begin{aligned}\mathcal{J}_g(\varphi) &= -\int_{(0,1)^2} g_1(x)J(\varphi(x), g_2(x))\,dx\\ &= -\int_{(0,1)^2}\Big\langle [g_1(x),0]\cdot \operatorname{cof} Dg(x), D\varphi(x)\Big\rangle\,dx,\end{aligned}$$

where the integration is with respect to the relevant components of the variation measure of g as earlier.

Let $\psi \in C^\infty_C[0,1)$ be such that $\psi \geq 0$, $\psi' \leq 0$ and

$$\int_{B(0,1)} \psi(|x|)\,dx = 1.$$

For each $\varepsilon > 0$ we denote by η_ε the usual convolution kernel, that is

$$\eta_\varepsilon(x) = \psi_\varepsilon(|x|) = \varepsilon^{-2}\psi\Big(\frac{|x|}{\varepsilon}\Big).$$

It is clear that $\eta_\varepsilon * D\varphi = D\eta_\varepsilon * \varphi$ converges uniformly to $D\varphi$ as $\varepsilon \to 0+$ and hence

$$\mathcal{J}_g(\varphi) = \lim_{\varepsilon\to 0_+} -\int_{(0,1)^2}\Big\langle [g_1(x),0]\cdot \operatorname{cof} Dg(x), \Big(\int_{B(x,\varepsilon)} \varphi D\eta_\varepsilon(x-z)\,dz\Big)\Big\rangle\,dx.$$

It is easy to see that $D\eta_\varepsilon(x) = {\psi_\varepsilon}'(|x|)\nu$, where $\nu = \frac{x}{|x|}$ is the normal vector. By Fubini theorem and change to polar coordinates we get

$$\mathcal{J}_g(\varphi) = -\lim_{\varepsilon\to 0_+}\int_{(0,1)^2}\varphi(z)\Big(\int_0^\varepsilon \psi'_\varepsilon(r)\int_{\partial B(z,r)} g_1(x)d(Dg|_{\partial B(z,r)}(x))dr\Big)dz.$$

By the degree formula Proposition 4.5 we obtain

$$\mathcal{J}_g(\varphi) = -\lim_{\varepsilon\to 0_+}\int_{(0,1)^2}\varphi(z)\Big(\int_0^\varepsilon \psi'_\varepsilon(r)\int_{\mathbb{R}^2} \deg(g,B(z,r),y)\,dy\,dr\Big)dz.$$

Let us fix $0 < \varepsilon < \frac{1}{2}\operatorname{dist}(\operatorname{supp}\varphi, \partial(0,1)^2)$. Then we have with the help of Lemma 6.1

$$\begin{aligned}|\mathcal{J}_g(\varphi)| &\leq 2\int_{\operatorname{supp}(\varphi)} |\varphi(z)|\Big(\int_0^\varepsilon |\psi'_\varepsilon(r)|\int_{\mathbb{R}^2} |\deg(g,B(z,r),y)|\,dy\,dr\Big)dz\\ &\leq 2\|\varphi\|_\infty\int_{(0,1)^2}\Big(\int_0^\varepsilon \frac{C}{\varepsilon^3}\int_{\mathbb{R}^2} N(g,B(z,r),y)\,dy\,dr\Big)dz \qquad (6.2)\\ &\leq C\|\varphi\|_\infty\int_{\mathbb{R}^2}\frac{1}{\varepsilon^2}\int_{(0,1)^2} N(g,B(z,\varepsilon),y)\,dz\,dy.\end{aligned}$$

Notice that for fixed $y \in \mathbb{R}^2$ we have

$$N(g,B(z,t),y) = \sum_{z_i\in g^{-1}(y)} \chi_{B(z_i,t)}(z).$$

With this we obtain from (6.2) that

$$|\mathcal{J}_g(\varphi)| \leq C\|\varphi\|_\infty\int_{\mathbb{R}^2} N(g,(0,1)^2,y)\,dy. \qquad (6.3)$$

By [24, Theorem 7.7] we see that

$$\int_{\mathbb{R}^2} N(g,(0,1)^2,y)\,dy \le \mathcal{H}^2(f_1^t((0,1)^2)).$$

Combining this with (6.3), it follows that for every $\varphi \in C_0^1((0,1)^2)$ we have

$$|\mathcal{J}_g(\varphi)| \le C\|\varphi\|_\infty \mathcal{H}^2\big(f_1^t((0,1)^2)\big) \tag{6.4}$$

with C independent of φ. By Hahn-Banach Theorem there is an extension to every $\varphi \in C_0((0,1)^2)$ which satisfies the same bound. By Riesz Representation Theorem there is a measure μ_t such that

$$\mathcal{J}_g(\varphi) = \int_{(0,1)^2} \varphi(x)\,d\mu_t(x) \text{ for every } \varphi \in C_0^1((0,1)^2).$$

By (6.4) and (6.1) we have

$$\int_0^1 \mu_t((0,1)^2)\,dt \le C\int_0^1 \mathcal{H}^2\big(f_1^t((0,1)^2)\big)\,dt < \infty$$

and thus $\mathcal{ADJ}\,Df \in \mathcal{M}((0,1)^3)$. □

6.1. **Dependence on the system of coordinates.** In principle the Definition 1.4 of $\mathcal{ADJ}\,Df \in \mathcal{M}$ depends on our coordinate system. Below we show that this notion is independent on the system of coordinates.

Corollary 6.2. *Let $\Omega \subset \mathbb{R}^3$ be a domain and $f \in BV(\Omega,\mathbb{R}^3)$ be a homeomorphism such that $\mathcal{ADJ}\,Df \in \mathcal{M}(\Omega,\mathbb{R}^{3\times3})$. Then $\mathcal{ADJ}\,Df \in \mathcal{M}(\Omega,\mathbb{R}^{3\times3})$ also for a different coordinate system.*

Proof. By Theorem 1.1 we know that $f^{-1} \in BV$. Hence $f \in BV_{\text{loc}}$ and $f^{-1} \in BV_{\text{loc}}$ and both of these do not depend on the choice of coordinate system. Thus by Theorem 1.2 we have $\mathcal{ADJ}\,Df \in \mathcal{M}(\Omega,\mathbb{R}^{3\times3})$ for any coordinate system. □

It is of course not true that the value of

$$|\mathcal{ADJ}\,Df|(\Omega)$$

is independent of coordinate system. In fact it might be more natural to define $|\mathcal{ADJ}\,Df|$ as an average over all directions (and not only 3 coordinate directions). Then, one could ask for the validity of (compare with (1.1))

$$|Df^{-1}|(f(\Omega)) = |\mathcal{ADJ}\,Df|(\Omega).$$

6.2. **The notion is stable under weak convergence.** For possible applications in the Calculus of Variations we need to know that the notion of distributional adjoint is stable under weak convergence.

Theorem 6.3. *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain. Let f_j, f be a BV homeomorphisms of $(0,1)^3$ onto Ω and assume that $f_j \to f$ uniformly and weak* in $BV((0,1)^3,\Omega)$. Further let $\mathcal{ADJ}\,Df_j \in \mathcal{M}((0,1)^3)$ and*

$$\sup_j |\mathcal{ADJ}\,Df_j|\big((0,1)^3\big) < \infty. \tag{6.5}$$

Then $\mathcal{ADJ}\,Df \in \mathcal{M}$.

Proof. By (6.5) and Theorem 1.1 we obtain that f_j^{-1} form a bounded sequence in $BV(\Omega,\mathbb{R}^3)$ and hence it has a weakly* converging subsequence. Thus we can assume (passing to a subsequence) that $f_j^{-1}\to h$ weakly* in BV and also strongly in L^1 (see [1, Corollary 3.49]). We define the pointwise representative of h as

$$h(y) := \limsup_{r\to 0}\frac{1}{|B(y,r)|}\int_{B(y,r)} h.$$

Now we need to show that $h=f^{-1}$. Fix $x_0\in(0,1)^3$ and $0<r<\operatorname{dist}(x_0,\partial(0,1)^3)$. We find $\delta>0$ so that $B(f(x_0),\delta)$ is compactly contained in $f(B(x_0,r))$. Since $f_j\to f$ uniformly we obtain that for j large enough we have

$$B(f(x_0),\delta)\subset f_j(B(x_0,r)).$$

It follows that

$$f_j^{-1}(B(f(x_0),\delta))\subset B(x_0,r) \text{ and hence } |h(f(x_0))-x_0|\le r$$

where we use that $f_j^{-1}\to h$ strongly in L^1 and that we have a proper representative of h. As the above inequality holds for every $r>0$ we obtain $h(f(x_0))=x_0$.

From $f\in BV$ and $f^{-1}=h\in BV$ we obtain $\mathcal{ADJ}\,Df\in\mathcal{M}((0,1)^3)$ by Theorem 1.2. □

Acknowledgments. The authors would like to thank Jan Malý for pointing their interest to Theorem 2.4 and for many valuable comments. The authors would also like to thank to Ulrich Menne for his information about literature on Lebesgue area.

References

[1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[2] J. Ball, Global invertibility of Sobolev functions and the interpenetration of matter, *Proc. Roy. Soc. Edinburgh Sect. A* **88**3–4 (1981), 315–328.

[3] L. Cesari, *Surface area*, Annals of Mathematics Studies, no. 35. Princeton University Press, Princeton, N. J., 1956. x+595 pp.

[4] P. G. Ciarlet and J Nečas, Injectivity and self-contact in nonlinear elasticity, *Arch. Rational Mech. Anal.* **97** 3 (1987), 171-188.

[5] S. Conti and C. De Lellis, Some remarks on the theory of elasticity for compressible Neohookean materials, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **2** (2003), 521–549.

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

[7] C. De Lellis, Some fine properties of currents and applications to distributional Jacobians, *Proc. Roy. Soc. Edinburgh* **132** (2002), 815–842.

[8] L. D'Onofrio and R. Schiattarella, On the total variations for the inverse of a BV-homeomorphism, *Adv. Calc. Var.* **6** (2013), 321–338.

[9] L. D'Onofrio, S. Hencl, J. Malý and R. Schiattarella, Note on Lusin (N) condition and the distributional determinant, *J. Math. Anal. Appl.* **439** (2016), 171–182.

[10] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag, New York.

[11] H. Federer, Hausdorff measure and Lebesgue area, *Proc. Nat. Acad. Sci. U. S. A.* **37** (1951), 90–94.

[12] I. Fonseca and W. Gangbo, *Degree Theory in Analysis and Applications*, Clarendon Press, Oxford, 1995.

[13] W. Gustin, Boxing inequalities, *J. Math. Mech.* **9** (1960), 229–239.
[14] A. Hatcher, *Algebraic Topology*, Campbridge University Press, Cambridge, 2002, xii+544pp.
[15] D. Henao and C. Mora-Corral, Invertibility and weak continuity of the determinant for the modelling of cavitation and fracture in nonlinear elasticity, *Arch. Rational Mech. Anal.* **197** (2010), 619–655.
[16] D. Henao and C. Mora-Corral, Fracture surfaces and the regularity of inverses for BV deformations, *Arch. Rational Mech. Anal.* **201** (2011), 575–629.
[17] D. Henao and C. Mora-Corral, Lusin's condition and the distributional determinant for deformations with finite energy, *Adv. Calc. Var.* **5** (2012), 355–409.
[18] S. Hencl and P. Koskela, Regularity of the inverse of a planar Sobolev homeomorphism, *Arch. Rational Mech. Anal* **180** (2006), 75–95.
[19] S. Hencl and P. Koskela, *Lectures on Mappings of finite distortion*, Lecture Notes in Mathematics 2096, Springer, 2014, 176pp.
[20] S. Hencl, Sharpness of the assumptions for the regularity of a homeomorphism, *Michigan Math. J.* **59** (2010), 667–678.
[21] S. Hencl, P. Koskela and J. Onninen, Homeomorphisms of bounded variation, *Arch. Rational Mech. Anal* **186** (2007), 351–360.
[22] T. Iwaniec, P. Koskela and J. Onninen, Mappings of finite distortion: Monotonicity and continuity, *Invent. Math.* **144** (2001), 507–531.
[23] T. Iwaniec and G. Martin, *Geometric function theory and nonlinear analysis*, Oxford Mathematical Monographs, Clarendon Press, Oxford 2001.
[24] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999.
[25] S. Müller, *Det* = *det* A remark on the distributional determinant, *C. R. Acad. Sci. Paris Sr. I Math.* **311** no. 1 (1990), 13–17.
[26] S. Müller and S. J. Spector, An existence theory for nonlinear elasticity that allows for cavitation, *Arch. Rational Mech. Anal.* **131,** no. 1 (1995), 1–66.
[27] S. Müller, S. J. Spector and Q. Tang, Invertibility and a topological property of Sobolev maps, *Siam J. Math. Anal.* **27** (1996), 959–976.

Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00 Prague 8, Czech Republic
E-mail address: hencl@karlin.mff.cuni.cz

Department of Mathematical Analysis, Sokolovska 83, Praha 8, 186 75, Charles University in Prague and Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193, Bellaterra (Barcelona), Spain
E-mail address: aapo.p.kauranen@gmail.com

Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00 Prague 8, Czech Republic
E-mail address: rami.luisto@gmail.com