AN IMPLICIT FUNCTION THEOREM FOR LIPSCHITZ MAPPINGS INTO METRIC SPACES

PIOTR HAJŁASZ AND SCOTT ZIMMERMAN

Abstract. We prove a version of the implicit function theorem for Lipschitz mappings $f : \mathbb{R}^{n+m} \supset A \to X$ into arbitrary metric spaces. As long as the pull-back of the Hausdorff content $\mathcal{H}^n_\infty$ by f has positive upper n-density on a set of positive Lebesgue measure, then, there is a local diffeomorphism G in $\mathbb{R}^{n+m}$ and a Lipschitz map $\pi : X \to \mathbb{R}^n$ such that $\pi \circ f \circ G^{-1}$, when restricted to a certain subset of A of positive measure, is the orthogonal projection of $\mathbb{R}^{n+m}$ onto the first n-coordinates. This may be seen as a qualitative version of a simlar result of Azzam and Schul [2]. The main tool in our proof is the metric change of variables introduced in [6].

In memoriam: William P. Ziemer (1934-2017)

1. Introduction

The classical implicit function theorem (IFT) ensures that the map is structurally very nice near points where the derivative of the map has a certain rank. In this paper, we present a version of the IFT for Lipschitz mappings $f : \mathbb{R}^{n+m} \supset A \to X$ into arbitrary metric spaces. It turns out that in the case of mappings into metric spaces, the upper density defined below will play a role of the Jacobian of f. For a measurable set $A \subset \mathbb{R}^k$, and $x \in A$, we define the *lower* and *upper n-densities* of a mapping $f : A \to X$ as

$$\Theta^{*n}(f,x) := \limsup_{r\to 0} \frac{\mathcal{H}^n_\infty(f(B(x,r)\cap A))}{\omega_n r^n}, \quad \Theta^n_*(f,x) := \liminf_{r\to 0} \frac{\mathcal{H}^n_\infty(f(B(x,r)\cap A))}{\omega_n r^n}.$$

These are simply the upper and the lower n-densities of the pull-back of $\mathcal{H}^n_\infty$ by f on A. Here ω_n is the volume of the unit ball in $\mathbb{R}^n$ and the $\mathcal{H}^n_\infty$ is the *Hausdorff content* defined for subsets of X by

$$\mathcal{H}^n_\infty(E) = \inf \frac{\omega_n}{2^n} \sum_{i=1}^{\infty} (\operatorname{diam} A_i)^n,$$

where the infimum is taken over all coverings of E, i.e. $E \subset \bigcup_{i=1}^\infty A_i$. Note that the Hausdorff content of *any* bounded set is finite, and, for an L-Lipschitz map $f : A \to X$, $\Theta^{*n}(f,x) \leq L^n$ for *all* $x \in A$.

The reader may want to compare these definitions with the definition (and properties) of the upper and lower densities of measures in [1, 12, 14].

2010 *Mathematics Subject Classification.* 53C23, 28A75, 54E40.

Key words and phrases. metric spaces; Lipschitz mappings; implicit function theorem.

P.H. was supported by NSF grant DMS-1800457.

The following observation will be useful throughout the paper:

$$\Theta^{*n}(f,x)=0 \quad \text{if and only if} \quad \lim_{d\to 0}\frac{\mathcal{H}^n_\infty(f(Q(x,d)\cap A))}{\omega_n d^n}=0 \tag{1.1}$$

where $Q(x,d)$ is the cube centered at x with side length d. (Here and in what follows, a cube has edges parallel to the coordinate axes.) The main result of the paper is as follows:

Theorem 1.1 (Metric IFT)**.** *Fix a metric space X, a set $A\subset\mathbb{R}^{n+m}$ with positive Lebesgue measure, and a Lipschitz mapping $f:A\to X$. Suppose $\Theta^{*n}(f,x)>0$ on a subset of A with positive Lebesgue measure. Then*

(A) $\mathcal{H}^n(f(A))>0$;
(B) *There is a set $K\subset A$ with positive Lebesgue measure, a bi-Lipschitz C^1-diffeomorphism $G:U\to G(U)\subset\mathbb{R}^{n+m}$ defined on an open set $U\supset K$ and a $\sqrt{n}$-Lipschitz map $\pi:X\to\mathbb{R}^n$ such that*

$$\pi\circ f\circ G^{-1}(x_1,\dots,x_n,y_1\dots,y_m)=(x_1,\dots,x_n)\quad \text{for all } (x,y)\in G(K)$$

is a projection on the first n coordinates when restricted to the set $G(K)$.

Moreover the mapping $F=f\circ G^{-1}$ defined on $G(K)$ satisfies

(C) $F^{-1}(F(x,y))\cap G(K)\subset\{x\}\times\mathbb{R}^m$ *for any $(x,y)\in G(K)$;*
(D) $F|_{(\mathbb{R}^n\times\{y\})\cap G(K)}$ *is bi-Lipschitz for any $y\in\mathbb{R}^m$.*

Remark 1.2. It follows from the proof that we can exhaust the set of points where $\Theta^{*n}(f,x)>0$ by sets K as in (B) up to a set of $\mathcal{H}^{n+m}$ measure zero. (See the application of Lemma 2.4 in the proof of Lemma 3.1 as well as Remark 3.5.)

Remark 1.3. The map $\pi:X\to\mathbb{R}^n$ is in fact 1-Lipschitz as a map from X to $(\mathbb{R}^n,\ell^\infty_n)$ where the norm ℓ^∞_n is defined by $\|(x_1,\dots,x_n)\|_\infty=\max_i|x_i|$. This will follow from our proof.

Remark 1.4. Statement (C) means that the preimage under F of any point in $F(G(K))=f(K)$ is contained in an m-dimensional subspace of $\mathbb{R}^{n+m}$ orthogonal to $\mathbb{R}^n$. For related results about the structure of preimages $f^{-1}(z)$ of Lipschitz maps, see [9, Theorem 1.2], [13, Theorem 4.16].

Remark 1.5. In fact, we will prove a quantitative lower bound in (D):

$$\|x_1-x_2\|_\infty\le d(F(x_1,y),F(x_2,y)) \tag{1.2}$$

for any $y\in\mathbb{R}^m$ and all $(x_1,y),(x_2,y)\in(\mathbb{R}^n\times\{y\})\cap G(K)$.

Remark 1.6. The classical implicit function theorem is stated using a condition about the rank of the derivative of f, and the condition $\Theta^{*n}(f,x)>0$ is a related one. Indeed, in the case $X=\mathbb{R}^n$, we will see in Proposition 5.2 that the Jacobian of f defined by $|J^nf|(x)=\sqrt{\det(Df)(Df)^T(x)}$ satisfies $\Theta^{*n}(f,x)=|J^nf|(x)$ almost everywhere. See also Lemma 3.3 for the case of mappings $f:A\to\ell^\infty$.

Remark 1.7. In the theorem we cannot replace the density condition $\Theta^{*n}(f,x)>0$ by the simpler measure condition $\mathcal{H}^n(f(A))>0$. Indeed, even in the Euclidean case, Kaufmann [10] constructed a surjective C^1 mapping $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$, $n\ge 2$, satisfying $\operatorname{rank}Df\le 1$ *everywhere*. For such a map, condition (B) cannot be satisfied since it would imply that $\operatorname{rank}Df\ge n$ on K.

Recall that a set $E \subset \mathbb{R}^{n+m}$ is *countably $\mathcal{H}^m$-rectifiable* if there are Lipschitz mappings $f_i : \mathbb{R}^m \supset E_i \to \mathbb{R}^{n+m}$, $i \in \mathbb{N}$, such that $\mathcal{H}^m(E \setminus \bigcup_{i=1}^{\infty} f(E_i)) = 0$. As a corollary of Theorem 1.1 we obtain

Corollary 1.8. *Fix a metric space X, a set $A \subset \mathbb{R}^{n+m}$ with positive Lebesgue measure, and a Lipschitz mapping $f : A \to X$. Suppose $\Theta^{*n}(f, \cdot) > 0$ almost everywhere in A. Then $f^{-1}(x)$ is countably $\mathcal{H}^m$-rectifiable for $\mathcal{H}^n$-almost all $x \in X$.*

See Section 4 for the proof. For related results see [9, Theorem 1.2], [13, Theorem 4.16].

Our result may be seen as a qualitative version of a theorem proven in 2012 by Azzam and Schul [2]. In that paper, the authors proved the following quantitative version of the IFT for Lipschitz mappings into metric spaces:

Theorem 1.9 (Quantitative metric IFT; Azzam and Schul, 2012)**.** *Fix a metric space X and a 1-Lipschitz mapping $f : \mathbb{R}^{n+m} \to X$. Suppose $0 < \mathcal{H}^n(f([0,1]^{n+m})) \leq 1$ and*

$$0 < \delta \leq \mathcal{H}^{n,m}_{\infty}(f, [0,1]^{n+m}) \tag{1.3}$$

for some $\delta > 0$. Then there are constants $\Lambda = \Lambda(n, m, \delta) > 1$ and $\eta = \eta(n, m, \delta) > 0$, a set $K \subset [0,1]^{n+m}$ with

$$\mathcal{H}^{n+m}(K) \geq \eta, \tag{1.4}$$

and a Λ-bi-Lipschitz homeomorphism $G : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ such that $F = f \circ G^{-1}$ satisfies

$$F^{-1}(F(x,y)) \cap G(K) \subset \{x\} \times \mathbb{R}^m \quad \text{for any } (x,y) \in G(K) \subset \mathbb{R}^{n+m}$$

and $F|_{(\mathbb{R}^n \times \{y\}) \cap G(K)}$ is Λ-bi-Lipschitz for any $y \in \mathbb{R}^m$.

The authors of [2] call $\mathcal{H}^{n,m}_{\infty}$ the *(n,m)-Hausdorff content of f*. It is defined for a Lipschitz map $f : Q \to X$ from a cube $Q \subset \mathbb{R}^{n+m}$ to a metric space by

$$\mathcal{H}^{n,m}_{\infty}(f, Q) = \inf \sum_{j=1}^{\infty} \mathcal{H}^n_{\infty}(f(Q_j)) d_j^m, \tag{1.5}$$

where the infimum is taken over all families of open pairwise disjoint cubes $Q_j \subset Q$ of side length d_j that cover Q up to a set of measure zero.

Note that Theorems 1.1 and 1.9 provide the same qualitative structure on the vertical and horizontal slices of the preimage of F. However, Theorem 1.9 is a *quantitative* version of the metric IFT in the sense that it provides the lower bound (1.4) which depends only on the dimensions m, n and δ from (1.3). Moreover, the mapping G is a *globally* defined C-bi-Lipschitz homomorphism where C depends only on m, n, and δ. Our result (Theorem 1.1) does not contain these quantitative conclusions. This is because the assumption (1.3) in Theorem 1.9 is much stronger than the assumption that $\Theta^{*n}(f, x) > 0$ on a set of positive measure. Indeed, Proposition 5.1 shows that the positivity of $\Theta^{*n}(f, x)$ follows from the assumption (1.3). In fact, for any $\varepsilon > 0$, one may construct a mapping $f : [0,1]^2 \to \mathbb{R}$ with $\Theta^{*1}(f, x) = 1$ almost everywhere so that the set $K \subset \mathbb{R}^2$ satisfying the conclusion of Theorem 1.1 (for a global bi-Lipschitz homeomorphism G) must satisfy $\mathcal{H}^2(K) < \varepsilon$ (and hence (1.4) cannot hold). See Proposition 5.3 for the construction and a detailed statement.

On the other hand, while the assumptions of Theorem 1.1 are much weaker than those of Theorem 1.9, some of the conclusions seem stronger: (1) As we already pointed out, the condition about positivity of $\Theta^{*n}(f,x)$ is much weaker than condition (1.3); (2) Azzam and Schul assume that $0 < \mathcal{H}^n(f([0,1]^{n+m})) \leq 1$ while we do not assume anything about the Hausdorff measure of the image. In fact, we prove the lower bound $\mathcal{H}^n(f(A)) > 0$ in (A) and finiteness of the measure of the image plays no role in our theorem; (3) Our mapping G is a bi-Lipschitz C^1 diffeomorphism while their mapping G is only a bi-Lipschitz map. However, their map is defined globally and ours is defined locally only; (4) While parts (C) and (D) are the same as the corresponding statements in Theorem 1.9, part (B) seems stronger than that. (C) and (D) easily follow from (B), but we do not know if (B) can be concluded from Theorem 1.9; (5) We obtain the quantitative lower bound estimate (1.2); (6) At last, but not least, our proof is *much* simpler than that in [2].

The classical IFT states that a C^1 mapping has a nice structure near a point where the derivative has rank of a certain order. However, the classical IFT does not provide *any* estimate for the size of the set where the map is nice. Our result has the same feature as the classical one: we do not obtain any estimate for the size of the set K except that it has a positive measure.

The main tool in the proof of Theorem 1.1 will be the metric change of variables introduced in [6]. This change of variables has been used to prove versions of Sard's theorem for Lipschitz mappings and BLD mappings into metric spaces [6, 7].

This paper is organized as follows. In Section 2 we collect basic definitions and lemmata needed in the proofs of Theorem 1.1 and Corollary 1.8. In Sections 3 and 4 we prove Theorem 1.1 and Corollary 1.8 respectively. Finally, in Section 5, we prove some other results that help us compare Theorems 1.1 and 1.9, we prove that the condition $\mathcal{H}^{n,m}_\infty(f,Q) > 0$ implies positivity of $\Theta^{*n}(f,x)$ on a set of positive measure (Proposition 5.1), we prove that, if $f : \mathbb{R}^{n+m} \supset A \to \mathbb{R}^n$ is Lipschitz, then $\Theta^{*n}(f,x) = \Theta^n_*(f,x) = |J^n f|(x)$ almost everywhere in A (Proposition 5.2), and we construct an example showing that we cannot obtain any lower bound for $\mathcal{H}^{n+m}(K)$ (Proposition 5.3).

Notation used in the paper is fairly standard. The n-dimensional Hausdorff measure will be denoted by $\mathcal{H}^n$. Note that in $\mathbb{R}^n$, $\mathcal{H}^n$ equals the Lebesgue measure and we will use Hausdorff measure notation in place of the Lebesgue measure. Occasionally we will write $|E|$ to denote the Lebesgue measure of E. Notation $\mathcal{H}^n_\infty$ will stand for the Hausdorff content defined above. The constant ω_n denotes the measure of the unit ball in $\mathbb{R}^n$. The Banach space of bounded *real valued* sequences will be denoted by ℓ^∞. Balls in metric spaces are denoted by $B(x,r)$, and $Q(x,d)$ denotes the Euclidean cube centered at x with side length d. All cubes are assumed to have edges parallel to the coordinate axes. Occasionally a k-dimensional ball in a Euclidean space will be denoted by $B^k(x,r)$. By a Λ-bi-Lipschitz homeomorphism $f : (X,d) \to (Y,\rho)$ we mean a homeomorphism satisfying $\Lambda^{-1}d(x,y) \leq \rho(f(x),f(y)) \leq \Lambda d(x,y)$. The tangent space to $\mathbb{R}^k$ at $x \in \mathbb{R}^k$ will be denoted by $T_x\mathbb{R}^k$. By C we will denote a general constant whose value may change in a single string of estimates. Writing $C = C(n,m)$, for example, indicates that the constant C depends on n and m only.

Acknowledgements. The authors would like to thank the referee for valuable comments that led to an improvement of the paper.

2. Preliminaries

In this section we collect basic definitions and results that will be used later on.

If $k > n$, then the $\mathcal{H}^n_\infty$ content of subsets of $\mathbb{R}^k$ is very different from their Hausdorff measure. For example $\mathcal{H}^n_\infty(E) < \infty$ for any bounded set $E \subset \mathbb{R}^k$, but $\mathcal{H}^n(B) = \infty$ for any k-ball $B \subset \mathbb{R}^k$. However, we have (see [14, Theorem 2.6])

Lemma 2.1. *$\mathcal{H}^n_\infty(E) = \mathcal{H}^n(E)$ for all sets $E \subset \mathbb{R}^n$.*

Lemma 2.2. *Every separable metric space admits an isometric embedding into ℓ^∞.*

Indeed, given $x_0 \in X$ and a dense set $\{x_i\}_{i=1}^\infty$ in a separable metric space (X, d),

$$X \ni x \mapsto \kappa(x) = (d(x, x_i) - d(x_i, x_0))_{i=1}^\infty \in \ell^\infty$$

is an isometric embedding. This is the well known Kuratowski embedding for metric spaces.

For a proof of the following elementary result, see [8, Corollary 4.1.7].

Lemma 2.3. *Let Y be a metric space, let $E \subset Y$ and let $f : E \to \ell^\infty$ be an L-Lipschitz mapping. Then there is an L-Lipschitz mapping $F : Y \to \ell^\infty$ such that $F|_E = f$.*

The idea of the proof is very simple. Each component f_i of f is L-Lipschitz and we define F by extending each of the components of f using the formula from the McShane extension. Then it is easy to verify that the resulting map is L-Lipschitz and it takes values in ℓ^∞.

Fix an integer $k \geq 1$, and suppose $A \subset \mathbb{R}^k$ is measurable. Recall that a function $f : A \to \mathbb{R}$ is *approximately differentiable* at $x \in A$ if there is a measurable set $A_x \subset A$ and a linear map $L : \mathbb{R}^n \to \mathbb{R}$ such that x is a density point of A_x and

$$\lim_{A_x \ni y \to x} \frac{|f(y) - f(x) - L(y - x)|}{|y - x|} = 0.$$

L is called the *approximate derivative* of f at x and is denoted by $ap\, Df(x)$. Recall also that $x \in E \subset \mathbb{R}^k$ is a *density point* of E if $\mathcal{H}^k(E \cap B(x, d))/(\omega_k d)^k \to 1$ as $d \to 0$.

If in addition $f : A \to \mathbb{R}$ is Lipschitz, then the approximate derivative $ap\, Df(x)$ exists for almost every $x \in A$. This follows from the McShane extension and Rademacher's theorem. Indeed, if $F : \mathbb{R}^k \to \mathbb{R}$ is a Lipschitz extension of f, then $\operatorname{ap} Df(x)$ exists at *all* points of the set

$$E = \{x \in A : \, x \text{ is a density point of } A \text{ and } F \text{ is differentiable at } x\}.$$

Moreover $\operatorname{ap} Df(x) = DF(x)$ at points of the set E.

For a Lipschitz map $f = (f_1, f_2, \dots) : A \to \ell^\infty$, we define the *component-wise approximate derivative* by

$$ap\, Df(x) := \begin{bmatrix} ap\, Df_1(x) \\ ap\, Df_2(x) \\ \vdots \end{bmatrix}$$

Since each component f_i is Lipschitz, $ap\, Df$ exists almost everywhere in A.

It is easy to see that the row and column ranks of this $\infty \times k$ matrix are equal, and $\operatorname{rank}(ap\, Df(x))$ equals the dimension of the image of $ap\, Df(x)$ in ℓ^∞. It follows in particular that $\operatorname{rank}(ap\, Df(x)) \leq k$.

Let V be a linear space of all real sequences. In particular, $\ell^\infty \subset V$, but we do not equip V with any norm or topology. If all components of a mapping $g = (g_1, g_2, \dots) : \mathbb{R}^k \to V$ are differentiable at a point x, we will say that g is *component-wise differentiable* at x and write

$$Dg(x) := \begin{bmatrix} Dg_1(x) \\ Dg_2(x) \\ \vdots \end{bmatrix}$$

We will also need the following result of Federer (for a proof, see [11, Theorem 1.69], [14, Theorem 5.3], [15]).

Lemma 2.4. *If $A \subset \mathbb{R}^k$ is measurable and $f : A \to \mathbb{R}$ is Lipschitz, then for any $\varepsilon > 0$ there is a function $g \in C^1(\mathbb{R}^k)$ such that*

$$\mathcal{H}^k(\{x \in A :\ f(x) \neq g(x)\}) < \varepsilon.$$

It is easy to see that if x_0 is a density point of the set

$$\{x \in A :\ f(x) = g(x)\}, \tag{2.1}$$

then $\operatorname{ap} Df(x_0)$ exists and $\operatorname{ap} Df(x_0) = Dg(x_0)$. In particular $Dg = \operatorname{ap} Df$ almost everywhere in the set (2.1).

The next lemma was proven in [6, Proposition 2.3].

Lemma 2.5. *Let $D \subset \mathbb{R}^k$ be a cube or ball, and let $f : D \to \ell^\infty$ be L-Lipschitz. Then*

$$\operatorname{diam}(f(D)) \leq C(k) L \mathcal{H}^k(D \setminus A)^{1/k},$$

where $A = \{x \in D :\ Df(x) = 0\}$ and Df is the component-wise derivative of f.

Finally, in the proof of Corollary 1.8 we will need

Lemma 2.6. *If $f : X \to Y$ is a Lipschitz mapping between metric spaces and $A \subset X$, $0 \leq m \leq n$, then*

$$\int_Y^* \mathcal{H}^{n-m}(f^{-1}(y) \cap A)\, d\mathcal{H}^m(y) \leq (\operatorname{Lip} f)^m \frac{\omega_{n-m}\omega_m}{\omega_n}\, \mathcal{H}^n(A).$$

Here $\int^*$ stands for the upper integral and $\operatorname{Lip} f$ is a Lipschitz constant of f. Federer [5, 2.10.25] proved this result under additional assumptions. The general case was obtained by Davies [3]. A detailed proof is given in [13, Theorem 2.4].

Corollary 2.7. *If $f : X \to Y$ is Lipschitz mapping between metric spaces and $A \subset X$, $\mathcal{H}^n(A) = 0$, $0 \le m \le n$, then $\mathcal{H}^{n-m}(f^{-1}(y) \cap A) = 0$ for $\mathcal{H}^m$ almost all $y \in Y$.*

3. Proof of Theorem 1.1

The proof is based on techniques developed in [6] (see also [7]). Consider a Lipschitz map $f : A \to \ell^\infty$ defined on a measurable set $A \subset \mathbb{R}^k$. Our first lemma shows that, if the rank of $\operatorname{ap} Df(x)$ is at least j on a set of positive measure, then, up to local diffeomorphisms, f fixes the first j coordinates on some non-null subset.

Lemma 3.1. *Suppose $f : A \to \ell^\infty$ is a Lipschitz map defined on a measurable set $A \subset \mathbb{R}^k$. If $\operatorname{rank}(\operatorname{ap} Df(x)) \ge j$ on a subset of A of positive $\mathcal{H}^k$-measure, then there is an open set $U \subset \mathbb{R}^k$, a set $K \subset A \cap U$ of positive $\mathcal{H}^k$-measure, a bi-Lipschitz C^1-diffeomorphism $G : U \to G(U) \subset \mathbb{R}^k$, and a permutation of a finite number of coordinates $\Psi : \ell^\infty \to \ell^\infty$ (which is an isometry of ℓ^∞) such that*

$$(\Psi \circ f \circ G^{-1})_i(x) = x_i \quad \text{for } i = 1, 2, \dots, j \text{ and } x \in G(K). \tag{3.1}$$

That is for $x \in G(K)$ we have

$$(\Psi \circ f \circ G^{-1})(x_1, \dots, x_n) = (x_1, \dots, x_j, (\Psi \circ f \circ G^{-1})_{j+1}(x), (\Psi \circ f \circ G^{-1})_{j+2}(x), \dots).$$

Proof. By restricting f to the set where $\operatorname{rank}(\operatorname{ap} Df(x)) \ge j$, we may assume that $\operatorname{rank}(\operatorname{ap} Df(x)) \ge j$ a.e. in A. Since $f = (f_1, f_2, \dots) : A \to \ell^\infty$ is Lipschitz, each component f_i of f is Lipschitz. Therefore, by applying Lemma 2.4 component-wise, we may choose $F \subset A$ with $\mathcal{H}^k(F) > 0$ and a mapping $g = (g_1, g_2, \dots) : \mathbb{R}^k \to V$ with $g_j \in C^1(\mathbb{R}^k)$ for every $j \in \mathbb{N}$ and such that $g = f$, $Dg = \operatorname{ap} Df$, and $\operatorname{rank} Dg = \operatorname{rank} \operatorname{ap} Df \ge j$ on F. Here, as before, V is the vector space consisting of all real valued sequences. (This is needed since sequences $(g_i(x))_{i=1}^\infty$ are not necessarily bounded.)

Lemma 3.2. *Fix $x_0 \in F$. Under the above assumptions, there is a bi-Lipschitz C^1-diffeomorphism $G : U \to G(U) \subset \mathbb{R}^k$ defined on a neighborhood U of x_0 and a permutation $\Psi : V \to V$ of a finite number of coordinates so that*

$$(\Psi \circ g \circ G^{-1})_i(x) = x_i \quad \text{for } i = 1, 2, \dots, j \text{ and } x \in G(U).$$

That is, $\Psi \circ g \circ G^{-1}$ fixes the first j coordinates on $G(U)$.

Proof. Since $\operatorname{rank} Dg(x_0) \ge j$, a certain $j \times j$ minor of $Dg(x_0)$ has rank j. By precomposing g with a permutation $\tilde{\Psi}$ of j variables in $\mathbb{R}^k$ and postcomposing it with a permutation Ψ of j variables in V, we have that

$$\tilde{g} = (\tilde{g}_1, \tilde{g}_2, \dots) = \Psi \circ g \circ \tilde{\Psi}$$

satisfies

$$\det \left[\frac{\partial \tilde{g}_m}{\partial x_\ell} (\tilde{\Psi}^{-1}(x_0)) \right]_{1 \le m, \ell \le j} \neq 0. \tag{3.2}$$

Let

$$H(x) = (\tilde{g}_1(x), \dots, \tilde{g}_j(x), x_{j+1}, \dots, x_k).$$

It follows from (3.2) that $\det DH(\tilde{\Psi}^{-1}(x_0)) \neq 0$, so H is a diffeomorphism in a neighborhood $\tilde{U}$ of $\tilde{\Psi}^{-1}(x_0)$. Replacing $\tilde{U}$ by a smaller open set, it follows that H is bi-Lipschitz. Now observe that

$$(\tilde{g} \circ H^{-1})_i(x) = x_i \quad \text{for } i = 1, 2, \ldots, j \text{ and } x \in H(\tilde{U}).$$

Therefore, if we write $G = H \circ \tilde{\Psi}^{-1}$, then $\Psi \circ g \circ G^{-1} = \tilde{g} \circ H^{-1}$ satisfies the claim of the lemma on the open set $U = \tilde{\Psi}(\tilde{U})$, U is a neighborhood of x_0, and $G(U) = H(\tilde{U})$. □

Now if x_0 is any density point of F, then the set $K = F \cap U$ has positive measure. Since $f = g$ on K, (3.1) follows because the permutation of coordinates $\Psi : V \to V$ maps $\ell^\infty \subset V$ to $\ell^\infty \subset V$ in an isometric way. This completes the proof of Lemma 3.1. □

Lemma 3.3. *Fix a measurable set $A \subset \mathbb{R}^k$ and $n \leq k$. Suppose $f : A \to \ell^\infty$ is a Lipschitz map. If $\Theta^{*n}(f, x) > 0$ on a subset of A of positive measure, then* $\operatorname{rank}(\operatorname{ap} Df(x)) \geq n$ *on a set of positive measure.*

Remark 3.4. Note that the above lemmata involve Lipschitz mappings into ℓ^∞. As we will see later, this will be sufficient in the setting of any metric space since the separable metric space $f(A)$ may be embedded isometrically into ℓ^∞ via the Kuratowski embedding.

Remark 3.5. In the following proof, we will see in particular that, for $j \in \{0, 1, 2, \ldots, n-1\}$, the set of points $x \in A$ where $\Theta^{*n}(f, x) > 0$, $\operatorname{ap} Df(x)$ exists, and $\operatorname{rank}(\operatorname{ap} Df(x)) = j$ must have measure zero.

Proof. Suppose to the contrary that $\Theta^{*n}(f, x) > 0$ on a set of positive measure and $\operatorname{rank}(\operatorname{ap} Df(x)) < n$ almost everywhere in A. Then there is $j \in \{0, 1, 2, \ldots, n-1\}$ and a set $F \subset A$ with $\mathcal{H}^k(F) > 0$ such that $\Theta^{*n}(f, x) > 0$ for all $x \in F$, $\operatorname{ap} Df(x)$ exists and $\operatorname{rank}(\operatorname{ap} Df(x)) = j$ for all $x \in F$.

According to Lemma 3.1, there is a permutation $\Psi : \ell^\infty \to \ell^\infty$ of a finite number of variables, an open set $U \subset \mathbb{R}^k$, a set $K \subset F \cap U$ with $\mathcal{H}^k(K) > 0$ and a bi-Lipschitz C^1-diffeomorphism $G : U \to G(U) \subset \mathbb{R}^k$ such that $\hat{f} = \Psi \circ f \circ G^{-1}$ defined on $\hat{A} = G(A \cap U)$ satisfies

$$\hat{f}_i(x) = x_i \quad \text{for } i = 1, 2, \ldots, j \text{ and } x \in \hat{K} \tag{3.3}$$

where $\hat{K} = G(K)$. Note that $\operatorname{ap} D\hat{f}(x)$ exists and $\operatorname{rank}(\operatorname{ap} D\hat{f})(x) = j$ for all $x \in \hat{K}$, because composition with a diffeomorphism and a permutation Ψ preserve approximate differentiability and the rank of the approximate derivative.

Assume that x_0 is a density point of K. Since $\Theta^{*n}(f, x) > 0$ for all $x \in K$, in order to arrive to a contradiction, it suffices to show that

$$\Theta^{*n}(f, x_0) = 0.$$

Note that $y_0 = G(x_0)$ is a density point of $\hat{K} = G(K)$ because diffeomorphisms map density points to density points.

The next lemma shows that it suffices to prove that

$$\Theta^{*n}(\hat{f}, y_0) = \limsup_{d \to 0} \frac{\mathcal{H}^n_\infty(\hat{f}(B(y_0, d) \cap \hat{A}))}{\omega_n d^n} = 0. \tag{3.4}$$

Lemma 3.6. *If $\Theta^{*n}(\hat{f}, y_0) = 0$, then $\Theta^{*n}(f, x_0) = 0$.*

Proof. Let $d > 0$ be so small that $B(y_0, d) \subset G(U)$. Since the diffeomorphism G^{-1} is bi-Lipschitz on $G(U)$, there is a constant $\Lambda > 0$ such that

$$B\left(x_0, \frac{d}{\Lambda}\right) \subset G^{-1}(B(y_0, d)).$$

Since the permutation of coordinates $\Psi : \ell^\infty \to \ell^\infty$ is an isometry, it follows that $\mathcal{H}^n_\infty(\hat{f}(E)) = \mathcal{H}^n_\infty(f(G^{-1}(E)))$ for any set E in the domain of $\hat{f}$. Therefore

$$\mathcal{H}^n_\infty(\hat{f}(B(y_0, d) \cap \hat{A})) = \mathcal{H}^n_\infty(f(G^{-1}(B(y_0, d) \cap \hat{A})) \geq \mathcal{H}^n_\infty(f(B(x_0, d/\Lambda) \cap A)),$$

so

$$\begin{aligned}\Theta^{*n}(\hat{f}, y_0) &= \limsup_{d \to 0} \frac{\mathcal{H}^n_\infty(\hat{f}(B(y_0, d) \cap \hat{A}))}{\omega_n d^n} \\ &\geq \Lambda^{-n} \limsup_{d \to 0} \frac{\mathcal{H}^n_\infty(f(B(x_0, d/\Lambda) \cap A))}{\omega_n (d/\Lambda)^n} = \Lambda^{-n} \Theta^{*n}(f, x_0)\end{aligned}$$

and the lemma follows. □

To conclude the proof of (3.4), we will apply the following lemma.

Lemma 3.7. *Assume $d > 0$ is such that $Q(y_0, d) \subset G(U)$ and*

$$\mathcal{H}^k(Q(y_0, d) \setminus \hat{K}) < \left(\frac{d}{M}\right)^k$$

for some positive integer M. Then $\hat{f}(Q(y_0, d) \cap \hat{A})$ can be covered by M^j balls of radius $CLdM^{-1}$ for some constant $C = C(k, n) > 0$, where L is the Lipschitz constant of $\hat{f}$. In particular, we have

$$\mathcal{H}^n_\infty(\hat{f}(Q(y_0, d) \cap \hat{A})) \leq \omega_n (CLd)^n M^{j-n}.$$

Before proving this lemma, we will see how it can be used to prove (3.4). Let $\varepsilon > 0$. Fix a positive integer M such that $(CL)^n M^{j-n} < \varepsilon$. (This is possible since $j - n < 0$.) Since y_0 is a density point of $\hat{K}$, there is $\delta > 0$ such that for $0 < d < \delta$, $Q(y_0, d) \subset G(U)$ satisfies

$$\mathcal{H}^k(Q(y_0, d) \setminus \hat{K}) < \frac{\mathcal{H}^k(Q(y_0, d))}{M^k} = \left(\frac{d}{M}\right)^k.$$

Hence, by Lemma 3.7, we have

$$\frac{\mathcal{H}^n_\infty(\hat{f}(Q(y_0, d) \cap \hat{A}))}{\omega_n d^n} \leq (CL)^n M^{j-n} < \varepsilon \quad \text{for } 0 < d < \delta$$

which, along with (1.1), implies that $\Theta^{*n}(\hat{f}, y_0) = 0$. That completes the proof of (3.4) once Lemma 3.7 has been verified. The proof of Lemma 3.7 is nearly identical to the proof of [6, Lemma 2.7], but we will include it here for completeness.

Proof of Lemma 3.7. Assume that a positive integer $M > 0$ and $d > 0$ satisfy $Q(y_0, d) \subset G(U)$ and

$$\mathcal{H}^k(Q(y_0, d) \setminus \hat{K}) < \left(\frac{d}{M}\right)^k.$$

Since the result is translation invariant, we may assume without loss of generality that

$$Q(y_0, d) = Q = [0, d]^j \times [0, d]^{k-j}.$$

According to Lemma 2.3, the L-Lipschitz mapping $\hat{f} : Q \cap \hat{A} \to \ell^\infty$ admits an L-Lipschitz extension $\tilde{f} : Q \to \ell^\infty$. According to Rademacher's theorem, $\tilde{f}$ is component-wise differentiable for almost all points in Q.

Divide $[0, d]^j$ into M^j cubes $\{Q_\nu\}_{\nu=1}^{M^j}$ with pairwise disjoint interiors each of edge length d/M. It suffices to show that each set

$$\hat{f}((Q_\nu \times [0, d]^{k-j}) \cap \hat{A}) \subset \tilde{f}(Q_\nu \times [0, d]^{k-j})$$

is contained in an ℓ^∞-ball of radius $CLdM^{-1}$ for some constant $C = C(k, n) > 0$. By our assumptions, for each ν we have

$$\mathcal{H}^k((Q_\nu \times [0, d]^{k-j}) \setminus \hat{K}) \leq \mathcal{H}^k(Q \setminus \hat{K}) < \left(\frac{d}{M}\right)^k.$$

Hence

$$\mathcal{H}^k((Q_\nu \times [0, d]^{k-j}) \cap \hat{K}) > \left(M^{-j} - M^{-k}\right) d^k.$$

According to Fubini's Theorem, we may therefore choose some $\rho \in Q_\nu$ such that

$$\mathcal{H}^{k-j}((\{\rho\} \times [0, d]^{k-j}) \cap \hat{K}) > \left(1 - M^{j-k}\right) d^{k-j}$$

and $\tilde{f}$ is component-wise differentiable at almost all points of $\{\rho\} \times [0, d]^{k-j}$. Hence

$$\mathcal{H}^{k-j}((\{\rho\} \times [0, d]^{k-j}) \setminus \hat{K}) < \left(\frac{d}{M}\right)^{k-j}. \tag{3.5}$$

According to (3.3), $\hat{f}$ fixes the first j coordinates in $\hat{K}$. Since $\hat{f} = \tilde{f}$ in $\hat{K}$ and $\operatorname{rank}(\operatorname{ap} D\hat{f}(x)) = j$ everywhere in $\hat{K}$, it follows that $\tilde{f}_i(x) = x_i$ for $i = 1, 2, \ldots, j$ and $x \in \hat{K}$ and $\operatorname{rank} D\tilde{f}(x) = j$ almost everywhere in $\hat{K}$. Therefore, the component-wise derivative of $\tilde{f}$ along $\{\rho\} \times [0, d]^{k-j}$ vanishes at almost all points in $(\{\rho\} \times [0, d]^{k-j}) \cap \hat{K}$. That is

$$D\big(\tilde{f}\big|_{\{\rho\} \times [0,d]^{k-j}}\big) = 0 \quad \text{a.e. in } (\{\rho\} \times [0, d]^{k-j}) \cap \hat{K}.$$

Therefore Lemma 2.5 applied to $\tilde{f} : \{\rho\} \times [0, d]^{k-j} \to \ell^\infty$ (with k replaced by $k-j$) together with (3.5) yield

$$\operatorname{diam}(\tilde{f}(\{\rho\} \times [0, d]^{k-j})) \leq CL\mathcal{H}^{k-j}((\{\rho\} \times [0, d]^{k-j}) \setminus \hat{K})^{1/(k-j)} \leq CLdM^{-1}.$$

Since the distance from any point in $Q_\nu \times [0, d]^{k-j}$ to the set $\{\rho\} \times [0, d]^{k-j}$ is at most $\operatorname{diam}(Q_\nu) = \sqrt{j} dM^{-1}$ and $\tilde{f}$ is L-Lipschitz, this implies that

$$\operatorname{diam}(\tilde{f}(Q_\nu \times [0, d]^{k-j})) \leq CLdM^{-1}$$

(for a larger value of C). This proves Lemma 3.7. □

This also completes the proof of (3.4) and hence that of Lemma 3.3. □

We now can finish the proof of Theorem 1.1.

Proof of Theorem 1.1. Since $f(A) \subset X$ is a separable metric space, there is an isometric embedding $\kappa : f(A) \to \ell^\infty$ (see Lemma 2.2). The mapping κ is 1-Lipschitz. According to Lemma 2.3, the map κ admits a 1-Lipschitz extension $\mathcal{K} : X \to \ell^\infty$.

Then $\bar{f} = \mathcal{K} \circ f = \kappa \circ f : A \to \ell^\infty$ is Lipchitz and $\Theta^{*n}(\bar{f}, x) > 0$ on a subset of A with positive measure (composition with an isometric map does not change the upper density).

It follows from Lemma 3.3 (with $k = n + m$) that $\operatorname{rank}\operatorname{ap} D\bar{f} \geq n$ on a set of positive measure. Therefore, according to Lemma 3.1, there is an open set $U \subset \mathbb{R}^{n+m}$, a subset $K \subset A \cap U$ with $\mathcal{H}^{n+m}(K) > 0$, a bi-Lipschitz C^1-diffeomorphism $G : U \to G(U) \subset \mathbb{R}^{n+m}$ and a permutation of finitely many coordinates $\Psi : \ell^\infty \to \ell^\infty$ such that

$$(\Psi \circ \bar{f} \circ G^{-1})_i(x) = x_i \quad \text{for } i = 1, 2, \dots, n \text{ and } x \in G(K). \tag{3.6}$$

Let

$$P : \ell^\infty \to \mathbb{R}^n, \quad P(x_1, x_2, \dots) = (x_1, x_2, \dots, x_n)$$

be the projection onto the first n coordinates. Then P is 1-Lipschitz as a mapping to $\mathbb{R}^n$ equipped with the ℓ^∞_n norm, $\|(x_1, \dots, x_n)\|_\infty = \max_i |x_i|$ and $\sqrt{n}$-Lipschitz as a mapping to $\mathbb{R}^n$ with the Euclidean metric. Therefore, it follows that the mapping

$$\pi : X \to \mathbb{R}^n, \quad \pi = P \circ \Psi \circ \mathcal{K}$$

is 1-Lipschitz as a mapping to $\mathbb{R}^n$ equipped with the norm ℓ^∞_n and $\sqrt{n}$-Lipschitz as a mapping to $\mathbb{R}^n$ with the Euclidean metric (see Remark 1.3).

If we swith to notation

$$(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m) := (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}),$$

then clearly, (3.6) means that $(\pi \circ f \circ G^{-1})(x, y) = x$ for $(x, y) \in G(K)$ which completes the proof of the statement (B).

To prove (A), suppose to the contrary that $\mathcal{H}^n(f(A)) = 0$. Then $\mathcal{H}^n(f(K)) = 0$ and hence

$$\mathcal{H}^n((\pi \circ f \circ G^{-1})(G(K)) = \mathcal{H}^n(\pi(f(K))) \leq (\sqrt{n})^n \mathcal{H}^n(f(K)) = 0, \tag{3.7}$$

because the $\sqrt{n}$-Lipschitz map π can increase the $\mathcal{H}^n$-measure no more than by a factor $(\sqrt{n})^n$. On the other hand, $G(K)$ has positive $\mathcal{H}^{n+m}$-measure so it follows from Fubini's theorem that its projection $(\pi \circ f \circ G^{-1})(G(K))$ onto the first n-coordinates has positive $\mathcal{H}^n$-measure which contradicts (3.7).

Parts (C) and (D) are easy consequences of part (B) as follows. Write $F = f \circ G^{-1}$.

Let $(x', y') \in F^{-1}(F(x, y)) \cap G(K)$. Then $F(x', y') = F(x, y)$ so $x' = \pi(F(x', y')) = \pi(F(x, y)) = x$ and hence $(x', y') = (x, y') \in \{x\} \times \mathbb{R}^m$ which proves (C).

To prove (D), fix $y \in \mathbb{R}^m$ and let $(x_1, y), (x_2, y) \in G(K)$. Let Λ be the Lipschitz constant of F on $G(K)$. Since $\pi : X \to (\mathbb{R}^n, \ell^\infty_n)$ is 1-Lipschitz we have

$$\begin{aligned} n^{-1/2}|x_1 - x_2| \leq \|x_1 - x_2\|_\infty &= \|\pi(F(x_1, y)) - \pi(F(x_2, y))\|_\infty \\ &\leq d(F(x_1, y), F(x_2, y)) \leq \Lambda |x_1 - x_2| \end{aligned}$$

which proves (D) along with the estimate (1.2). The proof is complete. □

4. Proof of Corollary 1.8

Since $\Theta^{*n}(f,x) > 0$ almost everywhere in A, we can exhaust A up to a set of $\mathcal{H}^{n+m}$ measure zero by a countable family of pairwise disjoint sets of positive $\mathcal{H}^{n+m}$ measure $\{K_i\}$, where each of the sets $K = K_i$ satisfies claim (B) of Theorem 1.1. Say $\{G_i\}$ are the associated bi-Lipschitz C^1-diffeomorphisms.

Let $W = \bigcup_{i=1}^{\infty} K_i$ and $Z = A \setminus W$ so $\mathcal{H}^{n+m}(Z) = 0$. Let $f_i = f|_{K_i}$ and let $F_i = f_i \circ G_i^{-1}$. Since F_i is defined on $G_i(K_i)$ only, we have from part (C) of Theorem 1.1 that for *any* $z \in X$, $F_i^{-1}(z)$ is contained in an m-dimensional affine subspace of $\mathbb{R}^{n+m}$ and hence $f_i^{-1}(z) = G_i^{-1}(F_i^{-1}(z))$ is contained in an m-dimensional submanifold (of class C^1). Therefore, for any $z \in X$,

$$f^{-1}(z) \cap W = \bigcup_{i=1}^{\infty} f_i^{-1}(z)$$

is countably $\mathcal{H}^m$-rectifiable as it is contained in a countable union of m-manifolds, and it remains to observe from Corollary 2.7 that $\mathcal{H}^m(f^{-1}(z) \setminus W) = \mathcal{H}^m(f^{-1}(z) \cap Z) = 0$ for $\mathcal{H}^n$ almost all $z \in X$. □

5. Comparing Theorems 1.1 and 1.9

Recall the (n,m)-Hausdorff content which was defined in (1.5). As mentioned in the introduction, the assumption that $\Theta^{*n}(f,x) > 0$ on a set of positive measure in Theorem 1.1 is weaker than the assumption of positive (n,m)-Hausdorff content of a cube in Theorem 1.9. We see this fact in the following proposition, the proof of which follows easily from the Vitali Covering Theorem.

Proposition 5.1. *Suppose $Q \subset \mathbb{R}^{n+m}$ is a cube, X is a metric space, and $f : Q \to X$ is Lipschitz. Then*

$$\mathcal{H}^{n,m}_{\infty}(f,Q) \leq \frac{\omega_n}{2^n}(n+m)^{n/2} \int_Q \Theta^n_*(f,x)\,dx \leq \frac{\omega_n}{2^n}(n+m)^{n/2} \int_Q \Theta^{*n}(f,x)\,dx.$$

Proof. In this proof $Q(x,d)$ and $\overline{Q}(x,d)$ will denote open and closed cubes in $\mathbb{R}^{n+m}$ respectively. Note that $Q(x,d) \subset B(x,\lambda d)$, where $\lambda = \frac{\sqrt{n+m}}{2}$.

The function $\Theta^n_*(f,\cdot)$ is integrable on Q since it is bounded. Fix $\varepsilon > 0$. Denote by A the set of all points in the interior of Q which are Lebesgue points of the function $\Theta^n_*(f,\cdot)$. Fix a point $x \in A$, and choose $d_x > 0$ small enough so that $Q(x,d_x) \subset B(x,\lambda d_x) \subset Q$. Choose a sequence $\{d^i_x\}_{i=1}^{\infty}$ with $d_x > d^i_x \searrow 0$ satisfying the following for each $d = d^i_x$:

$$\Theta^n_*(f,x) \leq \frac{1}{|Q(x,d)|} \int_{Q(x,d)} \Theta^n_*(f,y)\,dy + \frac{\varepsilon}{2}$$

and

$$\frac{\mathcal{H}^n_{\infty}(f(Q(x,d)))}{\omega_n(\lambda d)^n} \leq \frac{\mathcal{H}^n_{\infty}(f(B(x,\lambda d)))}{\omega_n(\lambda d)^n} \leq \Theta^n_*(f,x) + \frac{\varepsilon}{2}.$$

Both inequalities imply that

$$\mathcal{H}^n_\infty(f(Q(x,d)))d^m \le \omega_n\lambda^n\left(\int_{Q(x,d)}\Theta^n_*(f,y)\,dy+\varepsilon d^{n+m}\right)\quad\text{for all } x\in A \text{ and all } d=d^i_x.$$

The collection of closed cubes

$$\mathcal{Q}=\{\overline{Q}(x,d^i_x):\ x\in A,\ i\in\mathbb{N}\}$$

is a fine Vitali covering of A. Thus there is a countable, pairwise disjoint collection of cubes $\{\overline{Q}(x_j,d_j)\}$ in $\mathcal{Q}$ so that

$$\mathcal{H}^{n+m}\left(Q\setminus\bigcup_j Q(x_j,d_j)\right)=\mathcal{H}^{n+m}\left(A\setminus\bigcup_j\overline{Q}(x_j,d_j)\right)=0.$$

Since the cubes $Q(x_j,d_j)$ are open, pairwise disjoint, contained in Q, and they cover Q up to a set of measure zero, the definition of $\mathcal{H}^{n,m}_\infty(f,Q)$ yields

$$\begin{aligned}\mathcal{H}^{n,m}_\infty(f,Q)\le\sum_j\mathcal{H}^n_\infty(f(Q(x_j,d_j)))d_j^m&\le\omega_n\lambda^n\left(\sum_j\int_{Q(x_j,d_j)}\Theta^n_*(f,y)\,dy+\varepsilon\sum_j d_j^{n+m}\right)\\&=\omega_n\lambda^n\left(\int_Q\Theta^n_*(f,y)\,dy+\varepsilon|Q|\right).\end{aligned}$$

Sending $\varepsilon\to 0$ gives the desired result. □

The next result shows that $\Theta^n_*(f,x)$ is in fact equal to the Jacobian of f when f is a Lipschitz mapping to $\mathbb{R}^n$. This result is related to Lemma 3.3. Consider a mapping $f:\mathbb{R}^{n+m}\to\mathbb{R}^n$ which is differentiable at $x\in\mathbb{R}^{n+m}$. Define the Jacobian $|J^nf|(x)$ at x as follows:

$$|J^nf|(x)=\sqrt{\det(Df)(Df)^T(x)}.$$

Geometrically, it follows that, when $\operatorname{rank}Df(x)=n$, the Jacobian satisfies

$$|J^nf|(x)=\frac{\mathcal{H}^n(W_{x,r})}{\omega_n r^n}\quad\text{for any } r>0,\tag{5.1}$$

where

$$W_{x,r}=f(x)+Df(x)(B(0,r))\quad\text{for } B(0,r)\subset T_x\mathbb{R}^{n+m}\tag{5.2}$$

is the ellipsoid approximation (in $\mathbb{R}^n$) of $f(B(x,r))$. This Jacobian plays an important role in the so called co-area formula [16, Theorem 2.7.3].

Observe that if $\pi:T_x\mathbb{R}^{n+m}\to(\ker Df(x))^\perp\subset T_x\mathbb{R}^{n+m}$ is the orthogonal projection onto the n-dimensional subspace $(\ker Df(x))^\perp$, then $W_{x,r}=f(x)+Df(x)(\pi(B(0,r))$, so $W_{x,r}$ is (up to a translation by the vector $f(x)$) the image of the n-dimensional ball $\pi(B(0,r))\subset(\ker Df(x))^\perp$ of radius r under the linear map $Df(x)$. That is, $|J^nf|(x)$ is the ratio of the volume of the ellipsoid $W_{x,r}$ to the volume of $\pi(B(0,r))$.

If the rank of $Df(x)$ is less than n, we have $|J^nf|(x)=0$. Therefore $|J^nf|(x)>0$ if and only if $\operatorname{rank}Df(x)=n$. We similarly define the Jacobian of any Lipschitz mapping $f:\mathbb{R}^{n+m}\supset A\to\mathbb{R}^n$ using the approximate derivative.

Proposition 5.2. *Let $f : A \to \mathbb{R}^n$ be a Lipschitz map defined on a measurable set $A \subset \mathbb{R}^{n+m}$. Then*

$$\Theta^n_*(f,x) = \Theta^{*n}(f,x) = |J^n f|(x) \tag{5.3}$$

for almost every $x \in A$.

Note that combining this result with Proposition 5.1 gives the following for any cube $Q \subset \mathbb{R}^{n+m}$ and any Lipschitz $f : Q \to \mathbb{R}^n$:

$$\mathcal{H}^{n,m}_\infty(f,Q) \leq \frac{\omega_n}{2^n}(n+m)^{n/2} \int_Q |J^n f|(x)\, dx. \tag{5.4}$$

This inequality is essentially Lemma 6.13 in [2].

Proof. Assume first that $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ is an L-Lipschitz mapping defined on all of $\mathbb{R}^{n+m}$. It suffices to prove that (5.3) holds true at all points of differentiability of f.

Let $x \in \mathbb{R}^{n+m}$ be a point of differentiability of f. Given $L > \varepsilon > 0$, there is $\delta > 0$ such that

$$|f(y) - f(x) - Df(x)(y-x)| < \varepsilon r \quad \text{for all } 0 < r < \delta \text{ and } y \in B(x,r). \tag{5.5}$$

Assume first that $|J^n f|(x) = 0$. We will show that $\Theta^n_*(f,x) = \Theta^{*n}(f,x) = 0$.

Let $W_x = f(x) + Df(x)(T_x\mathbb{R}^{n+m})$ be an affine space through $f(x)$ (which is the image of the derivative in $\mathbb{R}^n$). Since $|J^n f|(x) = 0$, we have that $\dim W_x \leq n-1$ and hence

$$f(B(x,r)) \subset B(f(x), Lr) \cap \{z \in \mathbb{R}^n : \operatorname{dist}(z, W_x) < \varepsilon r\} \quad \text{for } 0 < r < \delta. \tag{5.6}$$

Since $\dim W_x = k \leq n-1$ we have that

$$\mathcal{H}^n_\infty(f(B(x,r))) \leq C(n)\varepsilon L^{n-1} r^n.$$

Indeed, the k-dimensional affine ball $B(f(x), Lr) \cap W_x \subset \mathbb{R}^n$ can be covered by

$$C\left(\frac{Lr}{\varepsilon r}\right)^k \leq C\left(\frac{L}{\varepsilon}\right)^{n-1}$$

balls in $\mathbb{R}^n$ of radius εr and centered at the points of $B(f(x), Lr) \cap W_x$. Then the balls with radii $2\varepsilon r$ and the same centers cover the set on the right hand side of (5.6), and hence they also cover $f(B(x,r))$. Since a ball of radius $2\varepsilon r$ has diameter $4\varepsilon r$ we have that

$$\mathcal{H}^n_\infty(f(B(x,r))) \leq \frac{\omega_n}{2^n}(4\varepsilon r)^n C\left(\frac{L}{\varepsilon}\right)^{n-1} = C(n)\omega_n \varepsilon r^n L^{n-1}.$$

Therefore,

$$\frac{\mathcal{H}^n_\infty(f(B(x,r)))}{\omega_n r^n} \leq C\varepsilon L^{n-1} \quad \text{for } 0 < r < \delta$$

which readily yields $\Theta^n_*(f,x) = \Theta^{*n}(f,x) = 0$.

Assume now that $|J^n f|(x) > 0$. Let $W_{x,r} = f(x) + Df(x)(B(0,r))$ be the ellipsoid considered in (5.2). Let $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the singular values of $Df(x)$ i.e., the lengths of the semiaxes of $W_{x,r}$ are $0 < \lambda_1 r \leq \lambda_2 r \leq \ldots \leq \lambda_n r$. ($\lambda_1 > 0$ because $|J^n f|(x) > 0$).

Consider the three concentric and homothetic ellipsoids (we further assume $0 < \varepsilon < \lambda_1$ so $1 - \varepsilon/\lambda_1 > 0$)

$$W_{x,(1-\varepsilon/\lambda_1)r} \subset W_{x,r} \subset W_{x,(1+\varepsilon/\lambda_1)r}.$$

The distance between the boundary of the ellipsoid $W_{x,r}$ and the boundaries of each of the other two ellipsoids equals εr since the distance between the homothetic ellipsoids is measured along the shortest semiaxes (as an easy exercise for the Lagrange multipliers). Therefore it follows from (5.5) that

$$W_{x,(1-\varepsilon/\lambda_1)r} \subset f(B(x,r)) \subset W_{x,(1+\varepsilon/\lambda_1)r} \quad \text{for } 0 < r < \delta. \tag{5.7}$$

Indeed, the right inclusion follows immediately from (5.5). The proof of the left inclusion is more intricate. Suppose to the contrary that

$$z \in W_{x,(1-\varepsilon/\lambda_1)r} \setminus f(B(x,r)).$$

Then using a 'radial' projection from z and estimate (5.5) one can construct a retraction of the ellipsoid $W_{x,r}$ to its boundary which is a contradiction. We leave details of a construction of a retraction to the reader.

It follows from Lemma 2.1 and (5.1) that for any $R > 0$

$$\mathcal{H}^n_\infty(W_{x,R}) = \mathcal{H}^n(W_{x,R}) = |J^n f|(x)\omega_n R^n$$

so (5.7) implies that for $0 < r < \delta$ we have

$$|J^n f|(x)\left(1 - \frac{\varepsilon}{\lambda_1}\right)^n \leq \frac{\mathcal{H}^n_\infty(f(B(x,r)))}{\omega_n r^n} \leq |J^n f|(x)\left(1 + \frac{\varepsilon}{\lambda_1}\right)^n$$

and letting $\varepsilon \to 0$ yields (5.3).

Note that the proof presented above is enough to establish (5.4).

We can now proceed to the proof of the result in the general case when $f : \mathbb{R}^{n+m} \supset A \to \mathbb{R}^n$ is Lipschitz.

Let $\tilde{f} : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be a Lipschitz extension of f. Assume that L is the Lipschitz constant of $\tilde{f}$. Note that $|J^n f| = |J^n \tilde{f}|$ at almost all points of A, and, by the proof presented above, $|J^n \tilde{f}|(x) = \Theta^n_*(\tilde{f}, x) = \Theta^{*n}(\tilde{f}, x)$ for almost all $x \in \mathbb{R}^{n+m}$. Note also that $\Theta^{*n}(\tilde{f}, x) \geq \Theta^{*n}(f, x)$, because in the case of $\Theta^{*n}(\tilde{f}, x)$ we consider the Hausdorff content of $\tilde{f}(B(x,r))$ while in the case of $\Theta^{*n}(f, x)$ we only consider the Hausdorff content of $f(B(x,r) \cap A) = \tilde{f}(B(x,r) \cap A)$.

Since for almost all $x \in A$ we have

$$|J^n f|(x) = |J^n \tilde{f}|(x) = \Theta^n_*(\tilde{f}, x) = \Theta^{*n}(\tilde{f}, x) \geq \Theta^{*n}(f, x) \geq \Theta^n_*(f, x),$$

it suffices to show that

$$\Theta^n_*(f, x) \geq |J^n \tilde{f}|(x) \quad \text{for almost all } x \in A. \tag{5.8}$$

For almost all $x \in A$ such that $|J^n f|(x) = 0$, this is particularly easy. Indeed, we have

$$\Theta^n_*(f, x) \geq 0 = |J^n f|(x) = |J^n \tilde{f}|(x),$$

so (5.8) is obvious.

We are left with the case when $|J^n f|(x) > 0$. Since we want to prove (5.8) almost everywhere, we can assume that x is a density point of A and $\tilde{f}$ is differentiable at x. Then $|J^n \tilde{f}|(x) = \Theta^{*n}(\tilde{f}, x) = \Theta^n_*(\tilde{f}, x)$, $\operatorname{ap} Df(x) = D\tilde{f}(x)$, and $|J^n \tilde{f}|(x) = |J^n f|(x) > 0$. In particular, we have $\operatorname{rank} D\tilde{f}(x) = n$.

The idea of the rest of the proof is simple. Since x is a density point of A, for small $r > 0$, the content $\mathcal{H}^n_\infty(f(B(x,r) \cap A)) = \mathcal{H}^n(\tilde{f}(B(x,r) \cap A))$ **is not much smaller** than $\mathcal{H}^n(\tilde{f}(B(x,r))) = \mathcal{H}^n_\infty(\tilde{f}(B(x,r)))$. Therefore dividing by $\omega_n r^n$ and passing to the liminf as $r \to 0$ gives

$$\liminf_{r\to 0} \frac{\mathcal{H}^n_\infty(f(B(x,r) \cap A))}{\omega_n r^n} + \varepsilon \geq \liminf_{r\to 0} \frac{\mathcal{H}^n_\infty(\tilde{f}(B(x,r)))}{\omega_n r^n} = \Theta^n_*(\tilde{f}, x) = |J^n \tilde{f}|(x)$$

for all $\varepsilon > 0$. Thus the main focus in the argument presented below is proving the phrase "**is not much smaller**". While the idea of the proof presented below is very geometric and relatively simple, the details are not.

By translating the coordinate system we may assume that $x = 0$. The ellipsoid $W_{0,r} = \tilde{f}(0) + D\tilde{f}(0)(B(0,r))$ is the image of the ball $B^{n+m}(0,r) \subset T_0\mathbb{R}^{n+m}$. By abusing notation we will identify the tangent space $T_0\mathbb{R}^{n+m}$ with $\mathbb{R}^{n+m}$. For example the same notation will be used for the ball $B^{n+m}(0,r)$ in the tangent space $T_0\mathbb{R}^{n+m}$, and for the ball $B^{n+m}(0,r) = 0 + B^{n+m}(0,r)$ in $\mathbb{R}^{n+m}$.

Since $\operatorname{rank} D\tilde{f}(0) = n$, we have $\dim \ker D\tilde{f}(0) = m$. Rotating the coordinate system in $\mathbb{R}^{n+m}$ we may assume that

$$\mathbb{R}^{n+m} = T_0\mathbb{R}^{n+m} = (\ker D\tilde{f}(0))^\perp \oplus (\ker D\tilde{f}(0)) = \mathbb{R}^n \oplus \mathbb{R}^m.$$

Let

$$\pi : \mathbb{R}^n \oplus \mathbb{R}^m \to \mathbb{R}^n \oplus \{0\} \subset \mathbb{R}^n \oplus \mathbb{R}^m$$

be the orthogonal projection. Note that the n-dimensional ball in the tangent space

$$B^n_0(r) := \pi(B^{n+m}(0,r)) = (\mathbb{R}^n \times \{0\}) \cap B^{n+m}(0,r) \subset T_0\mathbb{R}^{n+m}$$

has radius r and

$$W_{0,r} = \tilde{f}(0) + D\tilde{f}(0)(B^n_0(r)).$$

Let $\varepsilon > 0$ be given, then there is a positive integer M such that

$$|J^n \tilde{f}|(0) \left(1 - \frac{\sqrt{2}}{\lambda_1 M}\right)^n \left(1 - \frac{1}{2M}\right) - \frac{L^n}{2M} \geq |J^n \tilde{f}|(0) - \varepsilon, \tag{5.9}$$

where $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ are the singular values of $D\tilde{f}(0)$.

For any $r > 0$ and any $0 < t < 1$ let

$$V_{t,r} = (\mathbb{R}^n \times B^m(0, tr)) \cap B^{n+m}(0,r)$$

be the tr-cylinder around $B^n_0(r)$ inside of the ball $B^{n+m}(0,r)$. Clearly $\mathcal{H}^{n+m}(V_{t,r}) < \omega_n r^n \cdot \omega_m (tr)^m$ because $V_{t,r} \subset B^n(0,r) \times B^m(0,tr)$. Also, when t is small, the volume of $V_{t,r}$ must be close to the volume of this product of balls in the following sense:

$$\lim_{t\to 0} \frac{\mathcal{H}^{n+m}(V_{t,r})}{\omega_n r^n \cdot \omega_m (tr)^m} = 1.$$

Thus there is $0 < t_M < (1 - \frac{1}{2M})^{1/n}$ such that

$$\left(1 - \frac{1}{4M}\right) \omega_n \omega_m r^{n+m} t_M^m < \mathcal{H}^{n+m}(V_{t_M,r}) < \omega_n \omega_m r^{n+m} t_M^m. \tag{5.10}$$

Note that t_M depends on M but not on r because $V_{t,r} = rV_{t,1}$ (where $rE := \{rx : x \in E\}$ for $E \subset \mathbb{R}^{n+m}$).

Since $0 \in A$ is a density point of A, we may choose $\delta > 0$ depending on M so that for $0 < r < \delta$ we have

$$\mathcal{H}^{n+m}(V_{t_M,r} \setminus A) \leq \mathcal{H}^{n+m}(B^{n+m}(0,r) \setminus A) < \frac{1}{4M} \omega_n \omega_m r^{n+m} t_M^m \tag{5.11}$$

and hence

$$\mathcal{H}^{n+m}(V_{t_M,r} \cap A) = \mathcal{H}^{n+m}(V_{t_M,r}) - \mathcal{H}^{n+m}(V_{t_M,r} \setminus A) > \left(1 - \frac{1}{2M}\right) \omega_n \omega_m r^{n+m} t_M^m. \tag{5.12}$$

Since $\tilde{f}$ is differentiable at 0, we may also assume (by taking, if necessary, a smaller $\delta > 0$ depending on M) that

$$|\tilde{f}(y) - \tilde{f}(0) - D\tilde{f}(0)y| < \frac{r}{M} \quad \text{for all } 0 < r < \delta \text{ and } y \in B^{n+m}(0,r). \tag{5.13}$$

Let $0 < r < \delta$. For $b \in \mathbb{R}^m$ we define

$$B_b^n(r) = (\mathbb{R}^n \times \{b\}) \cap B^{n+m}(0,r).$$

If we regard $V_{t_M,r}$ as a cylinder with base $B^m(0, t_M r)$ (and with spherical caps), then the fibers (orthogonal to the base) are the n-balls $B_b^n(r)$ where b ranges over $B^m(0, t_M r)$.

We claim that the set of $b \in B^m(0, t_M r)$ which satisfy

$$\mathcal{H}^n(B_b^n(r) \cap A) > \left(1 - \frac{1}{2M}\right) \omega_n r^n \tag{5.14}$$

has positive $\mathcal{H}^m$-measure. Indeed, suppose to the contrary that

$$\mathcal{H}^n(B_b^n(r) \cap A) \leq \left(1 - \frac{1}{2M}\right) \omega_n r^n \quad \text{for } \mathcal{H}^m\text{-almost all } b \in B^m(0, t_M r).$$

Then it follows from Fubini's theorem that

$$\mathcal{H}^{n+m}(V_{t_M,r} \cap A) \leq \left(1 - \frac{1}{2M}\right) \omega_n r^n \cdot \omega_m (t_M r)^m$$

which contradicts (5.12). In other words, we have shown that the set of fibers of $V_{t_M,r}$ which see a "large" part of A has positive $\mathcal{H}^m$-measure.

Let $b \in B^m(0, t_M r)$ be such that (5.14) is satisfied. Then the radius R of the ball $B_b^n(r)$ satisfies

$$r \geq R > \left(1 - \frac{1}{2M}\right)^{1/n} r,$$

and since $D\tilde{f}(0)$ vanishes in the direction of $(0, b)$

$$\tilde{f}(0) + D\tilde{f}(0)(B_b^n(r)) = \tilde{f}(0) + D\tilde{f}(0)(\pi(B_b^n(r))) = \tilde{f}(0) + D\tilde{f}(0)(B_0^n(R)) = W_{0,R},$$

$$\mathcal{H}^n(W_{0,R}) = |J^n \tilde{f}|(0) \omega_n R^n.$$

Recall that

$$0 < t_M < \left(1 - \frac{1}{2M}\right)^{1/n} \quad \text{and} \quad b \in B^m(0, t_M r).$$

Therefore, $|b| < t_M r < R$. Thus by (5.13) and the Pythagorean theorem, we have

$$|\tilde{f}(y) - (\tilde{f}(0) + D\tilde{f}(0)y)| \leq M^{-1}\sqrt{(t_M r)^2 + R^2} < \sqrt{2}M^{-1}R \quad \text{for } y \in \partial B_b^n(r). \tag{5.15}$$

Since the distance between the boundaries of the ellipsoids (we assume that M is so large that $\sqrt{2}/M < \lambda_1$)

$$W_{0,(1-\sqrt{2}M^{-1}/\lambda_1)R} \subset W_{0,R}$$

equals $\sqrt{2}M^{-1}R$, it follows from (5.15) (as in (5.7)) that

$$W_{0,(1-\sqrt{2}M^{-1}/\lambda_1)R} \subset \tilde{f}(B_b^n(r)).$$

Therefore

$$\begin{aligned} \mathcal{H}^n(\tilde{f}(B_b^n(r))) &\geq \mathcal{H}^n(W_{0,(1-\sqrt{2}M^{-1}/\lambda_1)R}) = |J^n\tilde{f}|(0)\omega_n\left(1 - \frac{\sqrt{2}}{\lambda_1 M}\right)^n R^n \\ &> |J^n\tilde{f}|(0)\omega_n\left(1 - \frac{\sqrt{2}}{\lambda_1 M}\right)^n\left(1 - \frac{1}{2M}\right) r^n. \end{aligned} \tag{5.16}$$

Inequality (5.14) also implies that

$$\mathcal{H}^n(B_b^n(r) \setminus A) = \mathcal{H}^n(B_b^n(r)) - \mathcal{H}^n(B_b^n(r) \cap A) < \frac{1}{2M}\,\omega_n r^n.$$

Therefore

$$\mathcal{H}^n(\tilde{f}(B_b^n(r) \setminus A)) \leq \frac{L^n}{2M}\,\omega_n r^n. \tag{5.17}$$

We have

$$\tilde{f}(B_b^n(r)) = \tilde{f}(B_b^n(r) \cap A) \cup \tilde{f}(B_b^n(r) \setminus A) \tag{5.18}$$

so (5.18), (5.16), and (5.17) yield

$$\begin{aligned} \mathcal{H}^n(\tilde{f}(B_b^n(r) \cap A)) &\geq \mathcal{H}^n(\tilde{f}(B_b^n(r))) - \mathcal{H}^n(\tilde{f}(B_b^n(r) \setminus A)) \\ &\geq \omega_n r^n\left(|J^n\tilde{f}|(0)\left(1 - \frac{\sqrt{2}}{\lambda_1 M}\right)^n\left(1 - \frac{1}{2M}\right) - \frac{L^n}{2M}\right) \end{aligned}$$

and hence (5.9) yields

$$\begin{aligned} \frac{\mathcal{H}^n_\infty(f(B^{n+m}(0,r) \cap A))}{\omega_n r^n} &= \frac{\mathcal{H}^n(\tilde{f}(B^{n+m}(0,r) \cap A))}{\omega_n r^n} \geq \frac{\mathcal{H}^n(\tilde{f}(B_b^n(r) \cap A))}{\omega_n r^n} \\ &\geq |J^n\tilde{f}|(0)\left(1 - \frac{\sqrt{2}}{\lambda_1 M}\right)^n\left(1 - \frac{1}{2M}\right) - \frac{L^n}{2M} \\ &\geq |J^n\tilde{f}|(0) - \varepsilon \end{aligned}$$

for any $0 < r < \delta$. Therefore

$$\Theta^n_*(f, 0) = \liminf_{r \to 0} \frac{\mathcal{H}^n_\infty(f(B^{n+m}(0,r) \cap A))}{\omega_n r^n} \geq |J^n\tilde{f}|(0)$$

which completes the proof of (5.8) and hence that of Proposition 5.2. □

The following example provides evidence that, if the assumption (1.3) is replaced by the assumptions of Theorem 1.1, then the bound (1.4) and global bi-Lipschitz homeomorphism G cannot be recovered. In other words, even if the n-density of f satisfies $\Theta^{*n}(f,x) > 0$ on a set of positive measure, there is no universal constant $\eta > 0$ depending only on m, n, and δ so that $\mathcal{H}^{n+m}(K) > \eta$.

Proposition 5.3. *Fix a constant $\Lambda > 1$. For any $\varepsilon > 0$, there is a mapping $f : \mathbb{R}^{1+1} \supset [0,1]^2 \to \mathbb{R}$ with $\Theta^{*1}(f,x) = \Theta^1_*(f,x) = |J^1 f|(x) = 1$ a.e. satisfying the following: for any measurable set $K \subset [0,1]^2$ and any Λ-bi-Lipschitz homeomorphism $G : \mathbb{R}^2 \to \mathbb{R}^2$ such that $(f \circ G^{-1})|_{(\mathbb{R}\times\{y\})\cap G(K)}$ is Λ-bi-Lipschitz for any $y \in \mathbb{R}$, we have $\mathcal{H}^2(K) < \varepsilon$.*

Proof. Fix $\varepsilon > 0$ and choose $N \in \mathbb{N}$ large enough so that $\Lambda^4 2^{1-N}\sqrt{2} < \varepsilon$. For any $n \in \mathbb{N}$, define $f_n : [0, 2^{-(n-1)}]^2 \to [0, 2^{-n}]^2$ as follows:

$$f_n(x,y) = \begin{cases} (x,y) & \text{if } (x,y) \in [0,2^{-n}] \times [0,2^{-n}] \\ (2^{-(n-1)} - x, y) & \text{if } (x,y) \in [2^{-n}, 2^{-(n-1)}] \times [0,2^{-n}] \\ (x, 2^{-(n-1)} - y) & \text{if } (x,y) \in [0,2^{-n}] \times [2^{-n}, 2^{-(n-1)}] \\ (2^{-(n-1)} - x, 2^{-(n-1)} - y) & \text{if } (x,y) \in [2^{-n}, 2^{-(n-1)}] \times [2^{-n}, 2^{-(n-1)}] \end{cases}.$$

That is, we divide $[0, 2^{-(n-1)}]^2$ into four squares of equal size. On the lower left square, f_n is the identity mapping. On the upper left and lower right squares, f_n is a reflection over an edge onto the lower left square. On the upper right square, f_n is a reflection over both the bottom and left edges onto the lower left square.

Define $f : [0,1]^2 \to [0, 2^{-N}]$ to be a composition of N of these reflections together with the projection $\pi : \mathbb{R}^2 \to \mathbb{R}$ onto the first coordinate: $\pi(x,y) = x$. That is, we set

$$f := \pi \circ f_N \circ f_{N-1} \circ \cdots \circ f_2 \circ f_1$$

Clearly, f is Lipschitz.

Divide $[0,1]^2$ into $(2^N)^2$ squares $\{Q_i\}$ of side length 2^{-N}. Note that in each of the squares f is a composition of an isometry of $\mathbb{R}^2$ and the orthogonal projection to $\mathbb{R}$ so $|J^1 f| = 1$ and hence $\Theta^{*1}(f,x) = \Theta^1_*(f,x) = |J^1 f|(x) = 1$ a.e.

Let G be any Λ-bi-Lipschitz homeomorphism of $\mathbb{R}^2$ and $K \subset [0,1]^2$ be a measurable set such that $(f \circ G^{-1})|_{(\mathbb{R}\times\{y\})\cap G(K)}$ is Λ-bi-Lipschitz for any $y \in \mathbb{R}$. Write $F = f \circ G^{-1}$. For each $y \in \mathbb{R}$, we have

$$\mathcal{H}^1((\mathbb{R}\times\{y\}) \cap G(K)) \le \Lambda \mathcal{H}^1(F((\mathbb{R}\times\{y\}) \cap G(K))) \le \Lambda \mathcal{H}^1([0, 2^{-N}]) = \Lambda 2^{-N}.$$

Indeed, the first inequality is a consequence of the fact that $F|_{(\mathbb{R}\times\{y\})\cap G(K)}$ is Λ-bi-Lipschitz and the second inequality follows simply from the fact that the image of F is contained in $[0, 2^{-N}]$. Note also that $\operatorname{diam}(G(K)) \le \Lambda \operatorname{diam}(K) \le \Lambda\sqrt{2}$. In particular, $G(K)$ is contained in some square $Q = I_1 \times I_2$ where I_1 and I_2 are intervals of length $2\Lambda\sqrt{2}$. Thus

$$\begin{aligned} \mathcal{H}^2(K) \le \Lambda^2 \mathcal{H}^2(G(K)) = \Lambda^2 \int_Q \chi_{G(K)} &= \Lambda^2 \int_{I_2} \mathcal{H}^1((\mathbb{R}\times\{y\}) \cap G(K))\, dy \\ &\le \Lambda^2 \int_{I_2} \Lambda 2^{-N}\, dy = \Lambda^4 2^{1-N}\sqrt{2} < \varepsilon. \end{aligned}$$

□

References

[1] Ambrosio, L., Kirchheim, B.: Rectifiable sets in metric and Banach spaces. *Math. Ann.* 318 (2000), 527–555.

[2] Azzam, J., Schul, R.: Hard Sard: quantitative implicit function and extension theorems for Lipschitz maps. *Geom. Funct. Anal.* 22 (2012), 1062–1123.

[3] Davies, R. O.: Increasing sequences of sets and Hausdorff measure, *Proc. London Math. Soc.* 20 (1970), 222-236.

[4] Evans, L. C., Gariepy, R. F.: *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[5] Federer, H.: *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.

[6] Hajłasz, P., Malekzadeh, S.: On conditions for unrectifiability of a metric space. *Anal. Geom. Metr. Spaces* 3 (2015), 1–14.

[7] Hajłasz, P., Malekzadeh, S., Zimmerman, S.: Weak BLD mappings and Hausdorff measure. *Nonlinear Anal.* 177 (2018), 524-531.

[8] Heinonen, J., Koskela, P., Shanmugalingam, N., Tyson, J. T.: *Sobolev spaces on metric measure spaces. An approach based on upper gradients.* New Mathematical Monographs, 27. Cambridge University Press, Cambridge, 2015.

[9] Karmanova, M.: Rectifiable sets and coarea formula for metric-valued mappings. *J. Funct. Anal.* 254 (2008), 1410-1447.

[10] Kaufman, R.: A singular map of a cube onto a square. *J. Differential Geom.* 14 (1979), 593-594 (1981).

[11] Malý, J., Ziemer, W. P.: *Fine regularity of solutions of elliptic partial differential equations.* Mathematical Surveys and Monographs, 51. American Mathematical Society, Providence, RI, 1997.

[12] Mattila, P.: *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability.* Cambridge Studies in Advanced Mathematics, 44. Cambridge University Press, Cambridge, 1995.

[13] Reichel, L. P.: The coarea formula for metric space valued maps, Ph.D. thesis, ETH Zürich, 2009.

[14] Simon, L.: *Lectures on geometric measure theory.* Proceedings of the Centre for Mathematical Analysis, Australian National University, 3. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[15] Whitney, H.: On totally differentiable and smooth functions. *Pacific J. Math.* 1 (1951), 143-–159.

[16] Ziemer, W. P.: *Weakly differentiable functions. Sobolev spaces and functions of bounded variation.* Graduate Texts in Mathematics, 120. Springer-Verlag, New York, 1989.

P. Hajłasz: Department of Mathematics, University of Pittsburgh, 301 Thackeray Hall, Pittsburgh, PA 15260, USA, hajlasz@pitt.edu

S. Zimmerman: Department of Mathematics, University of Connecticut, 341 Mansfield Road U1009, Storrs, CT 06269, USA, scott.zimmerman@uconn.edu