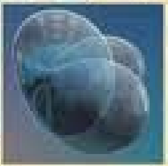
THIRD EDITION

GEOMETRIC MEASURE THEORY

A BEGINNER'S GUIDE



FRANK MORGAN

A BASIC INTRODUCTION OF GEOMETRIC MEASURE THEORY

QING HAN

Geometric measure theory studies properties of measures, functions and sets. In this note, we provide a basic introduction of geometric measure theory. We will study elementary properties of Hausdorff measures, Lipschitz functions and countably rectifiable sets. To make this note easily accessible, we confine our discussions in Euclidean spaces. Important topics include the area and coarea formulae for Lipschitz functions and approximate tangent space properties of countably rectifiable sets.

This note is prepared for lectures in the Special Lecture Series the author delivered in Peking University, June 2006. He would like to thank Professor Gang Tian for providing such an opportunity. He would also like to thank Miss Xianghui Yu for helping him dealing various issues when he was visiting Peking University. Last but not least, the author would like to thank all participants of this series of lectures. Because of their active participation in class and numerous discussions with him after the class, the author was able to correct many mistakes in the early version of this lecture note.

Address: Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556, USA.
Email: qhan@nd.edu.

1. Prerequisite Knowledge

This section serves as a preparation for the rest of the note. In the first part, we briefly review some basic theory of outer measures. The second part includes the covering theory. Many techniques in geometric measure theory involve covering arguments.

Let X be a topological space. We denote by $\mathcal{S}$ the family of all subsets of X. A collection $\mathcal{C}$ of subsets of X is said to be a *σ-algebra* if

(i) $\emptyset, X \in \mathcal{C}$;

(ii) $\cup_{i=1}^{\infty} A_i \in \mathcal{C}$ and $\cap_{i=1}^{\infty} A_i \in \mathcal{C}$ if $A_i \in \mathcal{C}$ for $i = 1, 2, 3, \cdots$;

(iii) $X \setminus A \in \mathcal{C}$ if $A \in \mathcal{C}$.

Let $\mathcal{B}$ be the smallest σ-algebra containing all open subsets of X. The elements of $\mathcal{B}$ are called *Borel subsets* of X. Obviously, $\mathcal{B}$ also contains all closed subsets of X.

Now we recall the definition of (outer) measures.

Let X be a topological space. If $\mu : \mathcal{S} \to [0, \infty]$ is a function such that $\mu(\emptyset) = 0$ and

$$\mu(A) \le \sum_{j=1}^{\infty} \mu(A_j),$$

for any $A \subset \cup_{j=1}^{\infty} A_j$, $A_j \subset \mathcal{S}$. Then μ is a *measure* on X.

A subset $A \subset X$ is μ-measurable if

$$\mu(B) = \mu(B \setminus A) + \mu(B \cap A) \quad \text{for any subset } B \subset X.$$

It is easy to show that the subfamily of $\mathcal{S}$ consisting of all μ-measurable subsets is a σ-algebra.

A measure μ on X is called a *Borel measure* if each Borel set is μ-measurable. A Borel measure μ is called *Borel regular* if for each subset $A \subset X$ there exists a Borel set $B \supset A$ such that $\mu(B) = \mu(A)$.

The following result is often referred to as the Caratheodory Criterion.

Theorem 1.1. *Let μ be a measure on a metric space X. Then all open sets are μ-measurable if and only if*

$$\mu(A \cup B) \ge \mu(A) + \mu(B),$$

for any subset $A \subset X$ and $B \subset X$ with dist$(A, B) > 0$.

Proof. The necessity is obvious. For the sufficiency, it suffices to prove that for any open subset $O \subset X$ and any subset $T \subset X$

$$\mu(T) \ge \mu(T \setminus O) + \mu(T \cap O),$$

if $\mu(T) < \infty$.

Set $C = T \cap O$, and for $k = 1, 2, \cdots$,

$$C_k = \left\{x \in C; \operatorname{dist}(x, X \setminus O) \ge \frac{1}{k}\right\}.$$

Then $\operatorname{dist}(C_k, T \setminus O) > 0$ and hence

$$\mu(T) \ge \mu(T \cap C_k) + \mu(T \setminus O).$$

Now we claim

$$\lim_{k\to\infty} \mu(T \cap C_k) = \mu(T \cap C) = \mu(T \cap O),$$

or $\lim_{k\to\infty} \mu(C_k) = \mu(C)$.

Let $C_0 = \emptyset$ and $R_k = C_k - C_{k-1}$ for $k \geq 1$. By the condition of the sufficiency, we obtain for any positive integer k

$$\mu(C_{2k}) \geq \sum_{i=1}^{k} \mu(R_{2i}),$$

$$\mu(C_{2k-1}) \geq \sum_{i=1}^{k} \mu(R_{2i-1}).$$

Note that

$$C = \bigcup_{k=1}^{\infty} C_k = C_{2k} + \bigcup_{i=k+1}^{\infty} R_{2i} + \bigcup_{i=k+1}^{\infty} R_{2i-1},$$

and

$$\sum_{i=1}^{\infty} \mu(R_{2i}) \leq \mu(T) < \infty,$$

$$\sum_{i=1}^{\infty} \mu(R_{2i-1}) \leq \mu(T) < \infty.$$

Hence for any positive integer k,

$$\mu(C) \leq \mu(C_{2k}) + \sum_{i=k+1}^{\infty} \mu(R_{2i}) + \sum_{i=k+1}^{\infty} \mu(R_{2i-1}).$$

This implies that $\lim_{k\to\infty} \mu(C_{2k}) = \mu(C)$. Similarly we have $\lim_{k\to\infty} \mu(C_{2k+1}) = \mu(C)$. □

Lemma 1.2. *Suppose that μ is a Borel regular measure on X and that*

$$X = \bigcup_{i=1}^{\infty} U_i,$$

where U_i is open and $\mu(U_i) < \infty$ for each $i = 1, 2, 3, \cdots$. Then

$$\mu(A) = \inf_{O \text{ open}, \, O \supset A} \mu(O),$$

for each subset $A \subset X$, and

$$\mu(A) = \sup_{C \text{ closed}, \, C \subset A} \mu(C),$$

for each μ-measurable subset $A \subset X$.

The proof is left as an exercise.

Suppose X is a locally compact and separable topological space. Then μ is a *Radon measure* if μ is Borel regular and finite on compact subsets of X.

Let U be a bounded open subset in $\mathbb{R}^n$. Denote by $\mathcal{M}(U)$ the space of signed Radon measures on U with finite mass and by $C_0(U)$ the space of continuous (real-valued) functions on U with compact support.

Definition 1.3. A sequence $\{\mu_k\}_{k=1}^{\infty} \subset \mathcal{M}(U)$ converges weakly to $\mu \in \mathcal{M}(U)$, denoted by $\mu_k \rightharpoonup \mu$ weakly in $\mathcal{M}(U)$, if

$$\int_U f d\mu_k \to \int_U f d\mu \quad \text{as } k \to \infty, \quad \text{for each } f \in C_0(U).$$

Lemma 1.4. *Assume that $\mu_k \rightharpoonup \mu$ weakly in $\mathcal{M}(U)$. Then*

$$\limsup_{k\to\infty} \mu_k(C) \le \mu(C),$$

for each compact set $C \subset U$, and

$$\mu(O) \le \liminf_{k\to\infty} \mu_k(O),$$

for each open set $O \subset U$.

Proof. (1) For any $\epsilon > 0$, there exist a $\delta > 0$ and a δ-neighborhood C_δ of C such that

$$\mu(C_\delta) \le \mu(C) + \epsilon.$$

By taking

$$f(x) = \min\{1, \frac{1}{\delta}\text{dist}(x, U \setminus C_\delta)\},$$

we have

$$\mu(C) + \epsilon > \int f d\mu = \lim_{k\to\infty} \int f d\mu_k \ge \limsup_{k\to\infty} \mu_k(C).$$

(2) For any $m < \mu(O)$, take a compact set $A \subset O$ with $\mu(A) > m$ and then take a positive number $\delta > 0$ such that the δ-neighborhood $A_\delta \subset O$. We define f similarly as above to obtain

$$m < \int f d\mu = \lim_{k\to\infty} \int f d\mu_k \le \liminf_{k\to\infty} \mu_k(O).$$

Obviously, this completes the proof. □

In the rest of the section, we discuss the Vitali Covering Lemma.

Theorem 1.5. *Suppose $\mathcal{B}$ is a family of closed balls in $\mathbb{R}^n$ with uniformly bounded radii. Then there is a pairwise disjoint subcollection $\mathcal{B}' \subset \mathcal{B}$ such that*

$$\bigcup_{B\in\mathcal{B}} B \subset \bigcup_{B'\in\mathcal{B}'} \hat{B}'.$$

Moreover, for any $B \in \mathcal{B}$, there exists an $S \in \mathcal{B}'$ such that $S \cap B \neq \emptyset$ and $B \subset \hat{S}$.

In Theorem 1.5, we use the notation $\hat{B} = \overline{B_{5r}(x)}$ if $B = \overline{B_r(x)}$. Obviously, $\mathcal{B}'$ is a countable collection.

Proof. Set

$$R = \sup\{\mathrm{rad}(B); B \in \mathcal{B}\} < \infty.$$

For $j = 1, 2, \cdots$, set

$$\mathcal{B}_j = \{B \in \mathcal{B}; R/2^j < \mathrm{rad}(B) \le R/2^{j-1}\}.$$

Then $\mathcal{B} = \cup_{j=1}^{\infty}\mathcal{B}_j$. We proceed to define $\mathcal{B}'_j \subset \mathcal{B}_j$ as follows:

(i) let $\mathcal{B}'_1$ be any maximal pairwise disjoint subcollection of $\mathcal{B}_1$;

(ii) assuming $j \ge 2$, and $\mathcal{B}'_1, \cdots, \mathcal{B}'_{j-1}$ are defined, let $\mathcal{B}'_j$ be a maximal pairwise disjoint subcollection of

$$\{B \in \mathcal{B}_j;\ B \cap B' = \emptyset \text{ whenever } B' \in \bigcup_{i=1}^{j-1}\mathcal{B}'_i\}.$$

Then evidently if $j \ge 1$ and $B \in \mathcal{B}_j$, we must have

$$B \cap B' \neq \emptyset \text{ for a } B' \in \bigcup_{i=1}^{j}\mathcal{B}'_i.$$

Otherwise, we contradict the maximality of $\mathcal{B}'_j$. For such a pair B and B', we have $\mathrm{rad}(B) \le R/2^{j-1} = 2R/2^j \le 2\mathrm{rad}(B')$, so that $B \subset \hat{B}'$.

To end the proof, we simply take $\mathcal{B}' = \cup_{i=1}^{\infty}\mathcal{B}'_i$. □

Theorem 1.5 is often put in the following form, which is easier to use.

Corollary 1.6. *Suppose A is a subset of $\mathbb{R}^n$ covered by a family of closed balls $\mathcal{B}$ in $\mathbb{R}^n$ with uniformly bounded radii. Then there is a pairwise disjoint subcollection $\{B_i\}_{i=1}^{\infty} \subset \mathcal{B}$ such that*

$$A \subset \bigcup_{i=1}^{\infty}\hat{B}_i.$$

A subset $A \subset \mathbb{R}^n$ is covered *finely* by a collection $\mathcal{B}$ of balls if, for any $\epsilon > 0$ and any $x \in A$, there is a ball $B \in \mathcal{B}$ such that $x \in B$ and $\mathrm{rad}(B) < \epsilon$.

Corollary 1.7. *Suppose A is a subset of $\mathbb{R}^n$ covered finely by a family of closed balls $\mathcal{B}$ in $\mathbb{R}^n$ with uniformly bounded radii. Then there is a pairwise disjoint subcollection $\mathcal{B}' \subset \mathcal{B}$ such that*

$$A \setminus \bigcup_{j=1}^{N} B_j \subset \bigcup_{B \in \mathcal{B}' \setminus \{B_1, \cdots, B_N\}} \hat{B},$$

for each finite subcollection $\{B_1, \cdots, B_N\} \subset \mathcal{B}'$.

Proof. Let $\mathcal{B}'$ be as in Theorem 1.5. For any $x \in A \setminus \cup_{j=1}^{N}B_j$, since $\mathcal{B}$ covers A finely and since $\mathbb{R}^n \setminus \cup_{j=1}^{N}B_j$ is open, we can find a $B \in \mathcal{B}$ with $B \cap (\cup_{j=1}^{N}B_j) = \emptyset$ and $x \in B$. By Theorem 1.5, there exists a $S \in \mathcal{B}'$ with $S \cap B \neq \emptyset$ and $B \subset \hat{S}$. Then $S \neq B_j$, $j = 1, \cdots, N$, and hence $x \in \hat{S}$. □

For a Borel measure μ on $\mathbb{R}^n$ and a subset $A \subset \mathbb{R}^n$, a family $\mathcal{B}$ of subset in $\mathbb{R}^n$ *covers μ-almost all* of A if $\mu(A \setminus \cup_{B \in \mathcal{B}} B) = 0$.

Corollary 1.8. *Suppose that A is a subset of $\mathbb{R}^n$ covered finely by a family of closed balls $\mathcal{B}$ in $\mathbb{R}^n$ with uniformly bounded radii and that μ is a Radon measure satisfying for a $\sigma \in (1, \infty)$*

$$\mu(\hat{B}) < \sigma\mu(B) \quad \text{for any } B \in \mathcal{B}.$$

Then, for each open $U \subset \mathbb{R}^n$, $\mathcal{B}$ contains a countable disjoint subfamily $\mathcal{C}$ that covers μ-almost all of $U \cap A$, with $\cup_{C\in\mathcal{C}} C \subset U$.

Proof. We first discuss the case where A is bounded. In this case, we also assume that U is bounded. Obviously, $\{B \in \mathcal{B}; B \subset U\}$ is a fine cover of $U \cap A$. By Corollary 1.7, there exists a disjoint subcollection $\mathcal{C} \subset \{B \in \mathcal{B}; B \subset U\}$ such that, for any finite subcollection $\{C_1, \cdots, C_N\}$,

$$A \setminus \bigcup_{i=1}^{N} C_i \subset \bigcup_{C\in\mathcal{C}\setminus\{C_1,\cdots,C_N\}} \hat{C}.$$

The collection $\mathcal{C}$ is countable because $\mu(U) < \infty$ and $\mu(C) > 0$ for $C \in \mathcal{C}$. We write $\mathcal{C} = \{C_i\}$. Note

$$\sum_{i=1}^{\infty} \mu(\hat{C}_i) \le \sigma \sum_{i=1}^{\infty} \mu(C) \le \sigma\mu(U) < \infty.$$

Now for any $N \ge 1$,

$$\mu(U \cap A \setminus \bigcup_{i=1}^{\infty} C_i) \le \mu(U \cap A \setminus \bigcup_{i=1}^{N} C_i) \le \sum_{i=N+1}^{\infty} \mu(\hat{C}_i).$$

We conclude the proof by letting $N \to \infty$.

Now we discuss the general case. We fix an $a \in X$. Starting with $U_0 = U$, $A_0 = \emptyset$ and $\mathcal{C}_0 = \{\emptyset\}$, we define open sets U_i and finite disjoint subfamilies $\mathcal{C}_i$ of $\mathcal{B}$ by induction so that

$$U_i = U_{i-1} \setminus \bigcup_{C\in\mathcal{C}_{i-1}} C, \quad A_i = \{x \in A; \operatorname{dist}(x, a) \le i\},$$

$$\bigcup_{C\in\mathcal{C}_i} C \subset U_i, \quad \mu(U_i \cap A_i \setminus \bigcup_{C\in\mathcal{C}_i} C) < \frac{1}{2^i}.$$

Then $\mathcal{C} = \cup_{i=1}^{\infty}\mathcal{C}_i$ is a countable disjoint subfamily of $\mathcal{B}$. Note $U_i = U \setminus \cup_{j=1}^{i-1} \cup_{C\in\mathcal{C}_j} C$. Hence

$$U \cap A \setminus \bigcup_{C\in\mathcal{C}} C \subset \bigcap_{k=1}^{\infty} \bigcup_{i=k}^{\infty} (U_i \cap A_i \setminus \bigcup_{C\in\mathcal{C}_i} C).$$

This implies for any $k \ge 1$

$$\mu(U \cap A \setminus \bigcup_{C\in\mathcal{C}} C) \le \sum_{i=k}^{\infty} \mu(U_i \cap A_i \setminus \bigcup_{C\in\mathcal{C}_i} C) < \frac{1}{2^{k-1}}.$$

We let $k \to \infty$. □

2. Hausdorff Measures

In this section, we introduce Hausdorff measures, a class of *lower dimensional* measures on $\mathbb{R}^n$, which allow us to measure certain *very small* subsets of $\mathbb{R}^n$.

Definition 2.1. Let $A \subseteq \mathbb{R}^n$ be a nonempty subset and $s \in [0, \infty)$.

(i) For any $\delta \in (0, \infty]$, define

$$\mathcal{H}^s_\delta(A) = \inf \big\{ \sum_{j=1}^{\infty} \omega(s) \big(\frac{\mathrm{diam} C_j}{2}\big)^s ;\ A \subset \bigcup_{j=1}^{\infty} C_j, \mathrm{diam} C_j \le \delta \big\},$$

where

$$\omega(s) = \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)}, \quad \text{for } s \in [0, \infty),$$

and

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx, \quad \text{for } s \in (0, \infty).$$

Note that $\Gamma(s)$ is the usual gamma function for $s \in (0, \infty)$.

(ii) Define

$$\mathcal{H}^s(A) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(A) = \sup_{\delta > 0} \mathcal{H}^s_\delta(A).$$

We call $\mathcal{H}^s$ the *s-dimensional Hausdorff measure* on $\mathbb{R}^n$.

For the empty set, we simply define $\mathcal{H}^s_\delta(\emptyset) = 0$ for any $\delta > 0$ and $\mathcal{H}^s(\emptyset) = 0$.

Note that $\omega(s)$ is the volume of the unit ball in $\mathbb{R}^s$ if s is a positive integer.

Remark 2.2. For $\delta > \delta'$, $\mathcal{H}^s_\delta(A) \le \mathcal{H}^s_{\delta'}(A)$. Hence $\mathcal{H}^s_\delta(\cdot)$ is a monotone decreasing function of $\delta \in (0, \infty]$. In particular, we have for any subset $A \subset \mathbb{R}^n$, $\delta > 0$ and $s \ge 0$

$$\mathcal{H}^s(A) \ge \mathcal{H}^s_\delta(A) \ge \mathcal{H}^s_\infty(A).$$

We note that it is necessary to require $\delta \to 0$ in order to force the coverings to follow the local geometry of the set A.

Observe that

$$\mathcal{L}^n\big(B_r(x)\big) = \omega(n) r^n,$$

for any balls $B_r(x) \subset \mathbb{R}^n$. Later on, we will prove that, if s is an integer, $\mathcal{H}^s$ agrees with the ordinary s-dimensional surface area on nice sets. This is the reason to include the normalizing constant $\omega(s)$ in the definition.

Theorem 2.3. *For any $s \in [0, \infty)$, $\mathcal{H}^s$ is a Borel regular measure on $\mathbb{R}^n$.*

Proof. (i) $\mathcal{H}^s$ is a measure, i.e., for any $\{A_i\}_{i=1}^\infty \subset \mathbb{R}^n$,

$$\mathcal{H}^s\big(\bigcup_{i=1}^\infty A_i\big) \le \sum_{i=1}^\infty \mathcal{H}^s(A_i).$$

By the definition of $\mathcal{H}^s_\delta$ for $\delta > 0$, we easily have

$$\mathcal{H}^s_\delta\big(\bigcup_{i=1}^\infty A_i\big) \le \sum_{i=1}^\infty \mathcal{H}^s_\delta(A_i).$$

Then take limit $\delta \to 0$.

(ii) $\mathcal{H}^s$ is a Borel measure. We will prove

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B),$$

for any Borel sets $A, B \subset \mathbb{R}^n$ with $\text{dist}(A, B) > 0$. It is easy to prove by the definition that

$$\mathcal{H}^s_\delta(A \cup B) = \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B),$$

if $\text{dist}(A, B) > 3\delta$. Then letting $\delta \to 0^+$, we obtain the desired result.

(iii) $\mathcal{H}^s$ is Borel regular. We need to show that, for any $A \subset \mathbb{R}^n$ with $\mathcal{H}^s(A) < \infty$, there exists a Borel set $B \supset A$ such that $\mathcal{H}^s(A) = \mathcal{H}^s(B)$.

Note that $\text{diam}\bar{C} = \text{diam}C$ for any subset $C \subset \mathbb{R}^n$. Hence,

$$\mathcal{H}^s_\delta(A) = \inf\Big\{\sum_{j=1}^\infty \omega(s)\big(\frac{\text{diam}C_j}{2}\big)^s; A \subset \bigcup_{j=1}^\infty C_j, \text{diam}C_j \leq \delta, C_j \text{ closed}\Big\}.$$

Since $\mathcal{H}^s(A) < \infty$, $\mathcal{H}^s_\delta(A) < \infty$ for any $\delta > 0$. For each $k \geq 1$, take closed sets $\{C_j^k\}_{j=1}^\infty$ so that $\text{diam}C_j^k \leq 1/k$, $A \subset \cup_{j=1}^\infty C_j^k$ and

$$\sum_{j=1}^\infty \omega(s)\big(\frac{\text{diam}C_j^k}{2}\big)^s \leq \mathcal{H}^s_{\frac{1}{k}}(A) + \frac{1}{k}.$$

Let $A_k = \cup_{j=1}^\infty C_j^k$ and $B = \cap_{j=1}^\infty A_k$. Then B is Borel and $A \subset B$, since $A \subset A_k$ for each k. Furthermore,

$$\mathcal{H}^s_{\frac{1}{k}}(B) \leq \sum_{j=1}^\infty \omega(s)(\frac{\text{diam}C_j^k}{2})^s \leq \mathcal{H}^s_{\frac{1}{k}}(A) + \frac{1}{k}.$$

Letting $k \to \infty$, we have $\mathcal{H}^s(B) \leq \mathcal{H}^s(A)$. With $A \subset B$, we get $\mathcal{H}^s(A) = \mathcal{H}^s(B)$. □

Theorem 2.4. *(i) $\mathcal{H}^0$ is the counting measure.*

(ii) $\mathcal{H}^s \equiv 0$ on $\mathbb{R}^n$ for any $s > n$.

(iii) $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ for any $\lambda > 0$.

(iv) $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$ for any affine isometries $L : \mathbb{R}^n \to \mathbb{R}^n$.

Proof. (iii) and (iv) are trivial. For (i), we note that $\omega(0) = 1$ and $\mathcal{H}^0(\{a\}) = 1$ for any $a \in \mathbb{R}^n$.

For (ii), we need to prove $\mathcal{H}^s(Q) = 0$ for any unit cube $Q \subset \mathbb{R}^n$. Fix an integer $m \geq 1$. Then the unit cube Q can be decomposed into m^n cubes with the side length $1/m$ and the diameter $\sqrt{n}/m$. Therefore,

$$\mathcal{H}^s_{\frac{\sqrt{n}}{m}}(Q) \leq \sum_{i=1}^{m^n} \omega(s)(\frac{\sqrt{n}}{m})^s = \omega(s) n^{s/2} m^{n-s},$$

and the last term goes to zero as $m \to \infty$ if $s > n$. Hence $\mathcal{H}^s(Q) = 0$, and $\mathcal{H}^s(\mathbb{R}^n) = 0$. □

The next result yields an easy method to prove that a set has a zero measure.

Lemma 2.5. *Suppose $A \subset \mathbb{R}^n$ and $\mathcal{H}^s_\delta(A) = 0$ for a $\delta \in (0, \infty]$. Then $\mathcal{H}^s(A) = 0$.*

Proof. We will show $\mathcal{H}^s(A) = 0$ if $\mathcal{H}^s_\infty(A) = 0$. It is obvious for $s = 0$. We consider $s > 0$.

For any $\epsilon > 0$, there exists a collection of sets $\{C_j\}_{j=1}^\infty$ such that $A \subset \cup_{j=1}^\infty C_j$ and

$$\sum_{j=1}^\infty \omega(s)\big(\frac{\mathrm{diam} C_j}{2}\big)^s \le \epsilon.$$

In particular, for each $j \ge 1$,

$$\mathrm{diam} C_j \le 2\big(\frac{\epsilon}{\omega(s)}\big)^{1/s} \equiv \delta(\epsilon).$$

Hence

$$\mathcal{H}^s_{\delta(\epsilon)}(A) \le \epsilon.$$

Since $\delta(\epsilon) \to 0$ as $\epsilon \to 0$, then $\mathcal{H}^s(A) = 0$. $\square$

The next result asserts that, for any subset $A \subset \mathbb{R}^n$, there is only one possible $s \ge 0$ such that $\mathcal{H}^s(A)$ is meaningful.

Lemma 2.6. *Let $A \subset \mathbb{R}^n$ and $0 \le s < t < \infty$.*

(i) If $\mathcal{H}^s(A) < \infty$, then $\mathcal{H}^t(A) = 0$.

(ii) If $\mathcal{H}^t(A) > 0$, then $\mathcal{H}^s(A) = \infty$.

Proof. We only need to prove (i), as (ii) follows from (i).

Consider $A \subset \mathbb{R}^n$ with $\mathcal{H}^s(A) < \infty$. For any $\delta > 0$, there exists a collection of sets $\{C_j\}_{j=1}^\infty$ such that $\mathrm{diam} C_j \le \delta$, $A \subset \cup_{j=1}^\infty C_j$ and

$$\sum_{j=1}^\infty \omega(s)\big(\frac{\mathrm{diam} C_j}{2}\big)^s \le \mathcal{H}^s_\delta(A) + 1 \le \mathcal{H}^s(A) + 1.$$

Then, we obtain

$$\begin{aligned}
\mathcal{H}^t_\delta(A) &\le \sum_{j=1}^\infty \omega(t)\big(\frac{\mathrm{diam} C_j}{2}\big)^t \\
&= \frac{\omega(t)}{\omega(s)} 2^{s-t} \sum_{j=1}^\infty \omega(s)\big(\frac{\mathrm{diam} C_j}{2}\big)^s (\mathrm{diam} C_j)^{t-s} \\
&\le \frac{\omega(t)}{\omega(s)} 2^{s-t} \delta^{t-s} \big(\mathcal{H}^s(A) + 1\big).
\end{aligned}$$

By letting $\delta \to 0$, we have $\mathcal{H}^t(A) = 0$. $\square$

Definition 2.7. The *Hausdorff dimension* of a set $A \subset \mathbb{R}^n$ is defined by

$$\dim_{\mathcal{H}}(A) = \inf\{s \in [0, \infty);\ \mathcal{H}^s(A) = 0\}.$$

Note that $\dim_{\mathcal{H}}(A) \le n$ by Theorem 2.4(ii). By setting $s = \dim_{\mathcal{H}}(A)$, we have $\mathcal{H}^t(A) = 0$ for any $t > s$ and $\mathcal{H}^t(A) = \infty$ for any $t < s$. We note that $\mathcal{H}^s(A)$ may be any number between 0 and ∞ inclusive. Furthermore, $\dim_{\mathcal{H}}(A)$ need not be an integer.

Even if $\dim_{\mathcal{H}}(A) = k$ is an integer and $0 < \mathcal{H}^k(A) < \infty$, A need not be a k-dimensional surface in any sense.

Example 2.8. The Cantor set in $\mathbb{R}$.

For any closed interval $J = [a, b]$ on $\mathbb{R}$ and any $t \in (2, \infty)$, we define

$$\Phi(J) = [a, a + \frac{b-a}{t}] \cup [b - \frac{b-a}{t}, b].$$

We have the following important identity

$$|J|^m = \sum_{S \in \Phi(J)} |S|^m \quad \text{for } m = \frac{\log 2}{\log t}.$$

In fact, the right side expression is $2\frac{|J|^m}{t^m}$. Hence $2 = t^m$. We begin with

$$H_0 = [0, 1],$$

and set inductively for $j = 1, 2, \cdots$,

$$H_j = \cup\{\Phi(J); J \in H_{j-1}\}.$$

Then we define the Cantor set by

$$C_t = \bigcap_{j=0}^{\infty} H_j.$$

It can be checked that

$$\dim_{\mathcal{H}}(C_t) = m = \frac{\log 2}{\log t},$$

and

$$\mathcal{H}^m(C_t) = \frac{\omega(m)}{2^m}.$$

Note that C_3 is the Cantor set studied in real analysis.

In the rest of this section, we prove that $\mathcal{H}^n$ agrees with $\mathcal{L}^n$ on $\mathbb{R}^n$. First we prove the following Isodiametric Inequality by the Steiner symmetrization.

Theorem 2.9. *For any $A \subset \mathbb{R}^n$,*

$$\mathcal{L}^n(A) \le \omega(n)\big(\frac{diam A}{2}\big)^n.$$

Proof. Without loss of generality, we assume A is closed and $\text{diam} A < \infty$. Taking a unit vector e in $\mathbb{R}^n$, we write

$$\mathbb{R}^n = \mathbb{R}^{n-1}_e \oplus \{e\},$$

where $\mathbb{R}^{n-1}_e$ is the orthogonal space of e in $\mathbb{R}^n$. Denote by $\Omega = P_e(A)$ the projection of A to $\mathbb{R}^{n-1}_e$. Then define

$$2h(x) = \mathcal{L}^1(A \cap L_x) \quad \text{for any } x \in \Omega,$$

where L_x is the line passing through x in the direction e. Set

$$S_e(A) = \{(x, y) \in \mathbb{R}^{n-1}_e \times \mathbb{R}; x \in \Omega, |y| \le h(x)\}.$$

It is easy to check by the Fubini Theorem that
(i) $h(x)$ is a measurable function;
(ii) $2\int_\Omega h(x)dx = \mathcal{L}^n(A)$.
Now we claim that diam$(S_e(A)) \le$diam(A). In fact, we prove

$$\begin{aligned}\operatorname{diam}\big(S_e(A)\big) &= \sup_{x,x'\in\Omega}\sqrt{(x-x')^2+\big(h(x)+h(x')\big)^2}\\ &\le \sup_{x,x'\in\Omega}\operatorname{diam}\big((L_x\cap A)\cup(L_{x'}\cap A)\big).\end{aligned}$$

Set

$$\begin{aligned}a &= \inf(L_x\cap A),\ \ b=\sup(L_x\cap A),\\ a' &= \inf(L_{x'}\cap A),\ \ b'=\sup(L_{x'}\cap A).\end{aligned}$$

Assume $b-a' \ge b'-a$ without loss of generality. Since $b'-a' \ge 2h(x')$ and $b-a \ge 2h(x)$ by the definition of h, we have

$$\begin{aligned}\operatorname{diam}\big((L_x\cap A)\cup(L_{x'}\cap A')\big) &\ge \sqrt{|x-x'|^2+|b-a'|^2}\\ &\ge \sqrt{|x-x'|^2+\big(h(x')+h(x)\big)^2}.\end{aligned}$$

This finishes the proof of the claim.

Next, take an orthogonal basis $\{e_1,\cdots,e_n\}$ of $\mathbb{R}^n$ and set

$$A^* = S_{e_n}\circ\cdots\circ S_{e_1}(A).$$

Then, diam$(A^*) \le$ diam(A) and $\mathcal{L}^n(A^*) = \mathcal{L}^n(A)$. It is clear that A^* is symmetric with respect to $e_1,\cdots,e_n$. Hence A^* is symmetric with respect to the origin. Thus $x\in A^*$ implies $-x\in A^*$. So for any $x\in A^*$, we have $2|x| \le \operatorname{diam}A^*$, or $x\in B_{\operatorname{diam}A^*/2}(0)$. This implies

$$A^* \subset B_{\operatorname{diam}A^*/2}(0).$$

Then

$$\mathcal{L}^n(A) = \mathcal{L}^n(A^*) \le \omega(n)\big(\frac{\operatorname{diam}A^*}{2}\big)^n \le \omega(n)\big(\frac{\operatorname{diam}A}{2}\big)^n.$$

This finishes the proof. □

Theorem 2.10. *$\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$.*

Proof. (i) We first show $\mathcal{L}^n(A) \le \mathcal{H}^n(A)$ for any $A\subset\mathbb{R}^n$. For any $\delta>0$, consider any collection of sets $\{C_j\}_{j=1}^\infty$ such that $A\subset\cup_{j=1}^\infty C_j$ and diam$C_j \le \delta$. Then by Theorem 2.9, we have

$$\mathcal{L}^n(A) \le \sum_{j=1}^{\infty}\mathcal{L}^n(C_j) \le \sum_{j=1}^{n}\omega(n)\big(\frac{\operatorname{diam}C_j}{2}\big)^n.$$

Taking the infimum, we have $\mathcal{L}^n(A) \le \mathcal{H}^n_\delta(A)$, and thus $\mathcal{L}^n(A) \le \mathcal{H}^n(A)$.

(ii) We prove that $\mathcal{H}^n$ is absolutely continuous with respect to $\mathcal{L}^n$. To this end, we first observe that for any $\delta > 0$ and $A \subset \mathbb{R}^n$,

$$\mathcal{H}^n_\delta(A) \leq \inf\{\sum_{i=1}^{\infty} \omega(n)\big(\frac{\text{diam}Q_j}{2}\big)^n; A \subset \bigcup_{j=1}^{\infty} Q_j,\ Q_j \text{ are cubes, } \text{diam}Q_j \leq \delta\}.$$

For each cube $Q \subset \mathbb{R}^n$, we have

$$\omega(n)\big(\frac{\text{diam}Q}{2}\big)^n = C(n)\mathcal{L}^n(Q),$$

where $C(n) = 2^{-n}\omega(n)\sqrt{n^n}$. Hence by the definition of Lebesgue measure,

$$\mathcal{H}^n_\delta(A) \leq C(n)\mathcal{L}^n(A),$$

or

$$\mathcal{H}^n(A) \leq C(n)\mathcal{L}^n(A).$$

(iii) We prove that $\mathcal{H}^n(A) \leq \mathcal{L}^n(A)$ for any $A \subset \mathbb{R}^n$. For any $\epsilon > 0$ and $\delta > 0$, there exists a collection of cubes $\{Q_j\}_{j=1}^\infty$ such that $\text{diam}Q_j < \delta$, $A \subset \cup_{j=1}^\infty Q_j$ and

$$\sum_{j=1}^{\infty} \mathcal{L}^n(Q_j) \leq \mathcal{L}^n(A) + \epsilon.$$

For each cube Q_i, choose disjoint balls $\{B^i_k\}_{k=1}^\infty \subset Q_i$ such that $\text{diam}B^i_k < \delta$ and $\mathcal{L}^n(Q_i \setminus \sum_{k=1}^\infty B^i_k) = 0$. Then by (ii), $\mathcal{H}^n(Q_i \setminus \sum_{k=1}^\infty B^i_k) = 0$. Hence we obtain

$$\begin{aligned}\mathcal{H}^n_\delta(A) &\leq \sum_{i=1}^{\infty} \mathcal{H}^n_\delta(Q_i) \leq \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \mathcal{H}^n_\delta(B^i_k) \leq \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \omega(n)\big(\frac{\text{diam}B^i_k}{2}\big)^n \\ &\leq \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \mathcal{L}^n(B^i_k) = \sum_{i=1}^{\infty} \mathcal{L}^n(Q_i) \leq \mathcal{L}^n(A) + \epsilon.\end{aligned}$$

This finishes the proof. □

3. Densities

In this section, we discuss densities for Hausdorff measures. We first define densities for general measures.

Definition 3.1. Let μ be a measure on a metric space X. For any subset $A \subset X$ and any point $x \in X$, the s-dimensional upper and lower densities of x in A in the measure μ, $\Theta^{*s}(\mu\lfloor A, x)$ and $\Theta^s_*(\mu\lfloor A, x)$, are defined, respectively, by

$$\Theta^{*s}(\mu\lfloor A, x) = \limsup_{r \to 0} \frac{1}{\omega(s)r^s}\mu\big(A \cap B_r(x)\big),$$

and

$$\Theta^s_*(\mu\lfloor A, x) = \liminf_{r \to 0} \frac{1}{\omega(s)r^s}\mu\big(A \cap B_r(x)\big).$$

The s-dimensional density $\Theta^s(\mu\lfloor A, x)$ is defined as the common value if $\Theta^{*s}(\mu\lfloor A, x) = \Theta^s_*(\mu\lfloor A, x)$.

It is easy to see that we may use closed balls in the above definition.

Remark 3.2. When $\mu = \mathcal{H}^s$, we write $\Theta^{*s}(A,x)$, $\Theta_*^s(A,x)$ and $\Theta^s(A,x)$ instead of $\Theta^{*s}(\mathcal{H}^s\lfloor A,x)$, $\Theta_*^s(\mathcal{H}^s\lfloor A,x)$ and $\Theta^s(\mathcal{H}^s\lfloor A,x)$. It is easy to show that $\Theta^{*s}(A,x)$ and $\Theta_*^s(A,x)$ are Borel functions of $x \in \mathbb{R}^n$ if $A \subset \mathbb{R}^n$ is Borel.

We first examine the densities of points in $\mathcal{L}^n$-measurable subsets of $\mathbb{R}^n$. By the Lebesgue differentiation theorem, we have for an $\mathcal{L}^n$-integral function f

$$\lim_{r\to 0} \frac{1}{\omega(n)r^n} \int_{B_r(x)} |f| = |f(x)|,$$

for $\mathcal{L}^n$-almost all $x \in \mathbb{R}^n$. For any $\mathcal{L}^n$-measurable subset A of $\mathbb{R}^n$, by taking $f = \chi_A$, we obtain

$$\lim_{r\to 0} \frac{1}{\omega(n)r^n} \mathcal{L}^n\big(A \cap B_r(x)\big) = \chi(x),$$

for $\mathcal{L}^n$-almost all $x \in \mathbb{R}^n$. Hence we conclude the following result.

Lemma 3.3. *Let A be an $\mathcal{L}^n$-measurable subset of $\mathbb{R}^n$. Then*

$$\Theta^n(A,x) = \begin{cases} 0 & \text{for } \mathcal{L}^n\text{-almost all } x \in \mathbb{R}^n \setminus A, \\ 1 & \text{for } \mathcal{L}^n\text{-almost all } x \in A. \end{cases}$$

Now we prove some results concerning densities of Hausdorff measures.

Theorem 3.4. *Suppose $E \subset \mathbb{R}^n$ is $\mathcal{H}^s$-measurable with $\mathcal{H}^s(E) < \infty$. Then*

$$\Theta^s(E,x) = 0 \quad \text{for } \mathcal{H}^s\text{-almost all } x \in \mathbb{R}^n \setminus E.$$

Proof. We assume that E is a Borel set. Fix a $t > 0$ and define

$$A_t = \Big\{x \in \mathbb{R}^n \setminus E; \limsup_{r\to 0} \frac{1}{\omega(s)r^s} \mathcal{H}^s\big(\overline{B_r(x)} \cap E\big) > t\Big\}.$$

We prove that $\mathcal{H}^s(A_t) = 0$ for each $t > 0$.

Note that $\mathcal{H}^s\lfloor E$ is a Radon measure. For any $\epsilon > 0$, there exists a compact set $K \subset E$ such that

$$\mathcal{H}^s(E \setminus K) \le \epsilon.$$

Consider the open set $U = \mathbb{R}^n \setminus K$. We have $A_t \subset U$. Fix a $\delta > 0$ and consider

$$\mathcal{B} = \Big\{B_r(x); B_r(x) \subset U, r \in (0,\delta), \frac{1}{\omega(s)r^s} \mathcal{H}^s\big(\overline{B_r(x)} \cap E\big) > t\Big\}.$$

Then $\mathcal{B}$ is a fine cover of A_t. By Corollary 1.7, there exists a countable disjoint family of balls $\{\overline{B_{r_i}(x_i)}\}_{i=1}^\infty$ in $\mathcal{B}$ such that

$$A_t \subset \bigcup_{i=1}^\infty \overline{B_{5r_i}(x_i)}.$$

Then, we have

$$\begin{aligned}\mathcal{H}^s_{10\delta}(A_t) &\le \sum_{i=1}^{\infty}\omega(s)(5r_i)^s \le \frac{5^s}{t}\sum_{i=1}^{\infty}\mathcal{H}^s(\overline{B_{r_i}(x_i)}\cap E)\\ &\le \frac{5^s}{t}\mathcal{H}^s(U\cap E) = \frac{5^s}{t}\mathcal{H}^s(E\setminus K) \le \frac{5^s}{t}\epsilon.\end{aligned}$$

Letting $\delta\to 0$, we have $\mathcal{H}^s(A_t)\le 5^s t^{-1}\epsilon$. Thus $\mathcal{H}^s(A_t)=0$ for any $t>0$. □

Theorem 3.5. *Suppose that $E\subset\mathbb{R}^n$ is $\mathcal{H}^s$-measurable with $\mathcal{H}^s(E)<\infty$. Then*

$$\Theta^{*s}(E,x)\le 1 \quad \textit{for } \mathcal{H}^s\textit{-almost all } x\in E.$$

Proof. We assume E is a Borel set. Fix $\epsilon>0$ and $t>1$, and define

$$E_t = \big\{x\in E; \limsup_{r\to 0}\frac{1}{\omega(s)r^s}\mathcal{H}^s\big(\overline{B_r(x)}\cap E\big) > t\big\}.$$

Since $\mathcal{H}^s\lfloor E$ is a Radon measure, there exists an open set U containing E_t with

$$\mathcal{H}^s(U\cap E)\le \mathcal{H}^s(E_t)+\epsilon.$$

Define

$$\mathcal{B} = \{B_r(x); B_r(x)\subset U, r\in(0,\delta),\ \frac{1}{\omega(s)r^s}\mathcal{H}^s\big(\overline{B_r(x)}\cap E\big)>t\}.$$

Then $\mathcal{B}$ is a fine cover of E_t. By Corollary 1.7, there exists a countable disjoint family of balls $\{\overline{B_{r_i}(x_i)}\}_{i=1}^{\infty}$ in $\mathcal{B}$ such that

$$E_t\subset \bigcup_{i=1}^{k}\overline{B_{r_i}(x_i)}\cup\bigcup_{i=k+1}^{\infty}\overline{B_{5r_i}(x_i)},$$

for each $k=1,2,\cdots$. Then

$$\begin{aligned}\mathcal{H}^s_{10\delta}(E_t) &\le \sum_{i=1}^{k}\omega(s)r_i^s + \sum_{i=k+1}^{\infty}\omega(s)(5r_i)^s\\ &\le \frac{1}{t}\sum_{i=1}^{k}\mathcal{H}^s(\overline{B_{r_i}(x_i)}\cap E) + \frac{5^s}{t}\sum_{i=k+1}^{\infty}\mathcal{H}^s(\overline{B_{r_i}(x_i)}\cap E)\\ &\le \frac{1}{t}\mathcal{H}^s(U\cap E) + \frac{5^s}{t}\mathcal{H}^s(\bigcup_{i=k+1}^{\infty}\overline{B_{r_i}(x_i)}\cap E),\end{aligned}$$

for any $k=1,2,\cdots$. By taking $k\to\infty$, we obtain

$$\mathcal{H}^s_{10\delta}(E_t)\le\frac{1}{t}\mathcal{H}^s(U\cap E)\le\frac{1}{t}\big(\mathcal{H}^s(E_t)+\epsilon\big).$$

Letting $\delta\to 0$ and then $\epsilon\to 0$, we have

$$\mathcal{H}^s(E_t)\le\frac{1}{t}\mathcal{H}^s(E_t).$$

Since $\mathcal{H}^s(E_t)\le\mathcal{H}^s(E)<\infty$, this implies $\mathcal{H}^s(E_t)=0$ for each $t>1$. □

Theorem 3.6. *Suppose that $E \subset \mathbb{R}^n$ is $\mathcal{H}^s$-measurable. Then*

$$\Theta^{*s}(\mathcal{H}^s_\infty \lfloor E, x) \geq \frac{1}{2^s} \quad \text{for } \mathcal{H}^s\text{-almost all } x \in E.$$

Proof. Fix a $t \in (0,1)$ and define

$$A_t = \big\{x \in E; \limsup_{r\to 0} \frac{1}{\omega(s)r^s}\mathcal{H}^s_\infty\big(\overline{B_r(x)} \cap E\big) < \frac{t}{2^s}\big\}.$$

We prove $\mathcal{H}^s(A_t) = 0$. For each positive integer j, define

$$B_j = \big\{x \in A_t; \frac{1}{\omega(s)r^s}\mathcal{H}^s_\infty\big(\overline{B_r(x)} \cap E\big) < \frac{t}{2^s} \text{ for any } r \in (0, \frac{1}{j})\big\}.$$

Then $B_j \subset B_{j+1}$ and $A_t = \cup_{j=1}^\infty B_j$. We prove $\mathcal{H}^s(B_j) = 0$ for each $j = 1, 2, \cdots$. For any $\delta < 1/j$, consider an arbitrary collection of subsets $\{C_i\} \subset \mathbb{R}^n$ such that $B_j \subset \cup_{i=1}^\infty C_i$, $B_j \cap C_i \neq \emptyset$ and $\text{diam}C_i < \delta$. Then for each i, there exists an $x \in B_j \cap C_i$ such that

$$\begin{aligned}\mathcal{H}^s_\delta(C_i \cap B_j) = \mathcal{H}^s_\infty(C_i \cap B_j) &\leq \mathcal{H}^s_\infty\big(B_{\text{diam}C_i}(x) \cap B_j\big)\\ &\leq \frac{t}{2^s}\omega(s)(\text{diam}C_i)^s = t\omega(s)\big(\frac{\text{diam}C_i}{2}\big)^s,\end{aligned}$$

where the first identity follows from the definition of Hausdorff measures and the fact that $\text{diam}C_i < \delta$. Then, we have

$$\mathcal{H}^s_\delta(B_j) \leq \sum_{i=1}^\infty \mathcal{H}^s_\delta(C_i \cap B_j) \leq t\omega(s)\sum_{i=1}^\infty\big(\frac{\text{diam}C_i}{2}\big)^s.$$

By taking the infimum of all such covers $\{C_i\}$ of B_j, we have

$$\mathcal{H}^s_\delta(B_j) \leq t\mathcal{H}^s_\delta(B_j).$$

This implies $\mathcal{H}^s_\delta(B_j) = 0$ for any $\delta < 1/j$, and hence $\mathcal{H}^s(B_j) = 0$. □

4. Lipschitz Functions

In this section, we discuss the differentiability of Lipschitz functions.

Definition 4.1. Let $A \subset \mathbb{R}^m$ be a subset.

(i) A function $f : A \to \mathbb{R}^n$ is Lipschitz if

$$|f(x) - f(y)| \leq C|x - y| \quad \text{for any } x, y \in A, \tag{4.1}$$

for a constant C. The smallest constant C such that (4.1) holds for any $x, y \in A$ is denoted by

$$\text{Lip}(f) = \sup\big\{\frac{|f(x) - f(y)|}{|x - y|}; x, y \in A, x \neq y\big\}.$$

(ii) A function $f : A \to \mathbb{R}^n$ is locally Lipschitz if f is Lipschitz in any compact subset $K \subset A$.

Our next result is fairly straightforward.

Lemma 4.2. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz, $A \subset \mathbb{R}^m$ and $0 \le s < \infty$. Then*

$$\mathcal{H}^s\big(f(A)\big) \le \big(Lip(f)\big)^s \mathcal{H}^s(A).$$

We will prove in the next section that $f(A)$ is $\mathcal{H}^m$-measurable in $\mathbb{R}^n$ if A is $\mathcal{L}^m$-measurable in $\mathbb{R}^m$.

Now, we discuss extensions of Lipschitz functions. We have the following result due to Kirszbraun.

Theorem 4.3. *If A is a subset in $\mathbb{R}^m$ and $f : A \to \mathbb{R}^n$ is Lipschitz, then there is a Lipschitz function $\hat{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\hat{f} = f$ on A and that $Lip(\hat{f})$ =$Lip(f)$.*

We first prove Theorem 4.3 for $n = 1$.

Proof for $n = 1$. Consider $f : A \to \mathbb{R}$. Define

$$\hat{f}(x) = \inf_{a \in A}\{f(a) + \mathrm{Lip} f|x - a|\} \quad \text{for any } x \in \mathbb{R}^m.$$

Since $f(a) + \mathrm{Lip} f \cdot |b - a| \ge f(b)$ for any $a, b \in A$, it follows that $\hat{f}$ is real-valued and $\hat{f}\,|_A = f$. Furthermore, for any $x, y \in \mathbb{R}^m$, we have

$$\begin{aligned}\hat{f}(x) &\le \inf_{a \in A}\{f(a) + \mathrm{Lip} f(|y - a| + |x - y|)\}\\ &= \hat{f}(y) + \mathrm{Lip} f|x - y|,\end{aligned}$$

i.e.,

$$\hat{f}(x) \le \hat{f}(y) + \mathrm{Lip} f|x - y|,$$

and similarly

$$\hat{f}(y) \le \hat{f}(x) + \mathrm{Lip} f|x - y|.$$

Hence, $|\hat{f}(x) - \hat{f}(y)| \le \mathrm{Lip} f|x - y|$ for any $x, y \in \mathbb{R}^m$. □

If we apply the result we just proved to the general case $f : A \to \mathbb{R}^n$, we get an extension $\hat{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\hat{f} = f$ on A and that $\mathrm{Lip}(\hat{f}) \le \sqrt{n}\mathrm{Lip}(f)$. This does not yield Theorem 4.3.

In order to prove Theorem 4.3, we need the following lemma.

Lemma 4.4. *Suppose P is a nonempty compact subset in $\mathbb{R}^m \times \{r \in \mathbb{R}; r \in (0, \infty)\}$ and, for any $t \in [0, \infty)$, set*

$$Y_t = \{y \in \mathbb{R}^m; y \in \overline{B_{rt}(a)} \text{ for any } (a, r) \in P\}.$$

Then $c = \inf\{t; Y_t \ne \emptyset\} < \infty$, Y_c consists of a single point b, and b belongs to the convex hull of

$$\Omega = \{a \in \mathbb{R}^m; (a, r) \in P \text{ and } b \in \partial B_{rc}(a) \text{ for some } r\}.$$

It is convenient here to identify P as a collection of closed balls $\{\overline{B_r(a)}\}$.

Proof. We note that each Y_t is compact and $0 \in Y_t$ for any $t \geq \sup\{|a|/r;\ (a,r) \in P\}$. We see that $Y_c = \cap\{Y_t; t \in (c,\infty)\}$ is not empty, as Y_t decreases with respect to t. We define

$$\mu = \sup\{r \in \mathbb{R}; (a,r) \in P \text{ for some } a\}.$$

If $y, z \in Y_c$, we then have for any $(a,r) \in P$

$$\begin{aligned}&\Big|\frac{1}{2}(y+z) - a\Big|^2 = \frac{1}{4}|y+z|^2 + |a|^2 - (y+z)\cdot a\\ =&\frac{1}{2}|y|^2 + \frac{1}{2}|z|^2 - \frac{1}{4}|y-z|^2 + |a|^2 - y\cdot a - z\cdot a\\ =&\frac{1}{2}\big(|y-a|^2 + |z-a|^2\big) - \frac{1}{4}|y-z|^2 \leq r^2c^2 - \frac{r^2}{4\mu^2}|y-z|^2.\end{aligned}$$

Thus $(y+z)/2 \in Y_t$ with $t = \sqrt{c^2 - |y-z|^2/(4\mu^2)}$. This implies $y = z$. We henceforth assume $Y_c = \{0\}$, by a translation in $\mathbb{R}^m$.

We take any $u \in \mathbb{R}^m$ with $|u| = 1$. Then $\varepsilon > 0$ implies $\varepsilon u \notin Y_c$. Hence there exists $(a,r) \in P$ with

$$|a|^2 \leq r^2c^2 < |\varepsilon u - a|^2 = \varepsilon^2 + |a|^2 - 2\varepsilon u \cdot a.$$

We note that (a,r) depends on ε. Since P is compact, then there exists an $(a,r) \in P$ such that $|a| = rc$ and $u \cdot a \leq 0$. Hence, there exists an $a \in \Omega$ such that $u \cdot a \leq 0$. Thus no $(m-1)$-dimensional plane separates 0 for the compact set Ω. □

Now we are ready to prove Theorem 4.3.

Proof of Theorem 4.3. We assume $\text{Lip}(f) = 1$ and consider the class Λ of all those Lipschitz extensions of f which map some subset of $\mathbb{R}^m$ into $\mathbb{R}^n$ and have the Lipschitz constant 1. By Hausdorff's maximal principle, Λ has a maximal (with respect to inclusion) element $g : T \to \mathbb{R}^m$, where T is a subset of $\mathbb{R}^m$.

We will show that if $p \in \mathbb{R}^m \setminus T$ there would exist $q \in \mathbb{R}^n$ with

$$|q - g(x)| \leq |p - x| \quad \text{whenever } x \in T.$$

Then we may extend g to $\tilde{g}$ by defining $\tilde{g}(p) = q$. Obviously, $\tilde{g} \in \Lambda$, and hence g would not be maximal in Λ. Thus we must prove that

$$\bigcap_{x\in T} \overline{B_{|x-p|}\big(g(x)\big)} \neq \emptyset.$$

Since these balls are compact, it suffices to verify that

$$\bigcap_{x\in F} \overline{B_{|x-p|}\big(g(x)\big)} \neq \emptyset,$$

for every finite subset F of T. For this purpose, we apply Lemma 4.4 with

$$P = \{\big(g(x), |x-p|\big); x \in F\}.$$

We choose distinct points $x_1, \cdots, x_k \in F$ and positive numbers $\lambda_1, \cdots, \lambda_k$ such that $g(x_i) \in \Omega$,

$$|b - g(x_i)| = |x_i - p|c \quad \text{for } i = 1, \cdots, k,$$

and

$$b=\sum_{i=1}^{k}\lambda_i g(x_i),\quad 1=\sum_{i=1}^{k}\lambda_i,$$

where Ω and b are as defined in Lemma 4.4. With the identity $2u\cdot v=|u|^2+|v|^2-|u-v|^2$, we obtain

$$\begin{aligned}0=&2\Big|\sum_{i=1}^{k}\lambda_i\big(g(x_i)-b\big)\Big|^2=2\sum_{i,j=1}^{k}\lambda_i\lambda_j\big(g(x_i)-b\big)\cdot\big(g(x_j)-b\big)\\ =&\sum_{i,j=1}^{k}\lambda_i\lambda_j\big(|g(x_i)-b|^2+|g(x_j)-b|^2-|g(x_i)-g(x_j)|^2\big)\\ \geq&\sum_{i,j=1}^{k}\lambda_i\lambda_j\big(c^2|x_i-p|^2+c^2|x_j-p|^2-|x_i-x_j|^2\big)\\ =&\sum_{i,j=1}^{k}\lambda_i\lambda_j\big(2c(x_i-p)\cdot c(x_j-p)+(c^2-1)|x_i-x_j|^2\big)\\ =&2\Big|c\sum_{i=1}^{k}\lambda_i(x_i-p)\Big|^2+(c^2-1)\sum_{i,j=1}^{k}\lambda_i\lambda_j|x_i-x_j|^2.\end{aligned}$$

Hence either $k=1$ and $c=0$ (because $p\neq x_1\in T$), or $k>1$ and $c\leq 1$. We conclude that $c\leq 1$, and hence $|g(x_i)-b|\leq|x_i-p|$ for any $i=1,\cdots,k$. □

Next, we prove a theorem of Rademacher concerning the differentiability of Lipschitz functions on Euclidean spaces.

Theorem 4.5. *Let $f:\mathbb{R}^m\to\mathbb{R}^n$ be a locally Lipschitz function. Then f is differentiable $\mathcal{L}^m$-almost everywhere in $\mathbb{R}^m$, i.e., for $\mathcal{L}^m$-almost every $x\in\mathbb{R}^m$, $Df(x)=(\partial_1 f(x),\cdots,\partial_m f(x))$ exists and*

$$\lim_{y\to x}\frac{1}{|y-x|}\big(f(y)-f(x)-Df(x)\cdot(y-x)\big)=0.$$

Proof. We assume $n=1$ and f is Lipschitz. Fix a vector $v\in\mathbb{R}^m$ with $|v|=1$, and define

$$\partial_v f(x)=\lim_{t\to 0}\frac{1}{t}\big(f(x+tv)-f(x)\big)\quad\text{for any } x\in\mathbb{R}^m,$$

if this limit exists. We define $\overline{\partial}_v f$ and $\underline{\partial}_v f$ similarly with lim replaced by lim sup and lim inf.

Claim 1. For each $v\in\mathbb{R}^m$ with $|v|=1$, $\partial_v f(x)$ exists for $\mathcal{L}^m$-almost all $x\in\mathbb{R}^m$.

To prove this, we first note that $\overline{\partial}_v f$ and $\underline{\partial}_v f$ are Borel functions, since, by the continuity of f,

$$\begin{aligned}\overline{\partial}_v f(x) &= \limsup_{t\to 0} \frac{1}{t}\big(f(x+tv)-f(x)\big)\\ &= \lim_{k\to\infty} \sup_{-\frac{1}{k}<t<\frac{1}{k}, t \text{ rational}} \frac{1}{t}\big(f(x+tv)-f(x)\big)\end{aligned}$$

is a Borel function, and so is

$$\underline{\partial}_v f(x) = \liminf_{t\to 0} \frac{1}{t}\big(f(x+tv)-f(x)\big).$$

Thus the set

$$\begin{aligned}B_v &= \{x\in\mathbb{R}^m; \partial_v f(x) \text{ does not exist}\}\\ &= \{x\in\mathbb{R}^m; \underline{\partial}_v f(x) < \overline{\partial}_v f(x)\}\end{aligned}$$

is Borel measurable.

Now for each $x, v\in\mathbb{R}^m$, with $|v|=1$, define $\phi:\mathbb{R}\to\mathbb{R}$ by $\phi(t)=f(x+tv)$. Then ϕ is Lipschitz, thus absolutely continuous. Hence ϕ is differentiable $\mathcal{L}^1$-almost everywhere. This implies

$$\mathcal{H}^1(B_v\cap L)=0,$$

for each line L parallel to v. Then the Fubini Theorem yields

$$\mathcal{L}^m(B_v)=0.$$

As a consequence, we see that

$$Df(x)\equiv\big(\partial_1 f(x),\cdots,\partial_m f(x)\big)$$

exists for $\mathcal{L}^m$-almost all $x\in\mathbb{R}^m$. This finishes the proof of Claim 1.

Claim 2. For each $v\in\mathbb{R}^m$ with $|v|=1$, $\partial_v f(x) = v\cdot Df(x)$ for $\mathcal{L}^m$-almost all $x\in\mathbb{R}^m$.

To see this, we take an arbitrary test function $\xi\in C_0^\infty(\mathbb{R}^m)$. Then

$$\int_{\mathbb{R}^m}\frac{1}{t}\big(f(x+tv)-f(x)\big)\xi(x)dx = -\int_{\mathbb{R}^m} f(x)\cdot\frac{1}{t}\big(\xi(x)-\xi(x-tv)\big)dx.$$

Let $t=\dfrac{1}{k}$ in the above equality for $k=1,2,\cdots$, and note that

$$\Big|k\big(f(x+\frac{1}{k}v)-f(x)\big)\Big| \le \mathrm{Lip}(f)|v| = \mathrm{Lip}f.$$

Thus the Dominated Convergence Theorem implies

$$\begin{aligned}&\int_{\mathbb{R}^m}\partial_v f(x)\xi(x)dx = -\int_{\mathbb{R}^m} f(x)\partial_v\xi(x)dx\\ &= -\sum_{i=1}^m v_i\int_{\mathbb{R}^m} f(x)\partial_i\xi(x)dx = \sum_{i=1}^m v_i\int_{\mathbb{R}^m}\partial_i f(x)\xi(x)dx\\ &= \int_{\mathbb{R}^m} v\cdot Df(x)\xi(x)dx,\end{aligned}$$

where we used the Fubini theorem and the absolute continuity of f on lines. This implies Claim 2.

Now choose a countable dense subset $\{v_k\}_{k=1}^{\infty}$ of $\mathbb{S}^{m-1}$. Set

$$A_k = \{x \in \mathbb{R}^m; \partial_{v_k} f(x), Df(x) \text{ exist and } \partial_{v_k} f(x) = v_k \cdot Df(x)\},$$

and define $A = \cap_{k=1}^{\infty} A_k$. Then $\mathcal{L}^m(\mathbb{R}^m \setminus A) = 0$, $\partial_{v_k} f(x) = v_k \cdot Df(x)$ for any $x \in A$ and $k = 1, 2, \cdots$.

Claim 3. f is differentiable at each point $x \in A$.

Fix an $x \in A$. For any $v \in \mathbb{S}^{m-1}$ and $t \in \mathbb{R}$ with $t \neq 0$, define

$$Q(x, v, t) = \frac{1}{t}\big(f(x+tv) - f(x)\big) - v \cdot Df(x).$$

Then for any $v' \in \mathbb{S}^{m-1}$, we have

$$\begin{aligned}
&|Q(x,v,t) - Q(x,v',t)| \\
\leq &\Big|\frac{1}{t}\big(f(x+tv) - f(x+tv')\big)\Big| + |(v-v') \cdot Df(x)| \\
\leq &\mathrm{Lip}(f)|v-v'| + |Df||v-v'| \\
\leq &(\sqrt{m}+1)\mathrm{Lip} f |v-v'|.
\end{aligned}$$

By the compactness of $\mathbb{S}^{m-1}$, for any $\epsilon > 0$, choose N large enough so that, for any $v \in \mathbb{S}^{m-1}$, there exists a $k \in \{1, 2, \cdots, N\}$ satisfying

$$|v - v_k| \leq \frac{\epsilon}{2(\sqrt{m}+1)\mathrm{Lip} f}.$$

Hence

$$\lim_{t \to 0} Q(x, v_k, t) = 0 \quad \text{for any } k = 1, 2, \cdots, N,$$

and there exists a $\delta > 0$ so that

$$|Q(x, v_k, t)| < \frac{\epsilon}{2} \quad \text{for any } 0 < |t| < \delta \text{ and } k = 1, 2, \cdots, N.$$

Thus for each $v \in \mathbb{S}^{m-1}$, there exists a $k \in \{1, 2, \cdots, N\}$ such that

$$|Q(x,v,t)| \leq |Q(x,v_k,t)| + |Q(x,v,t) - Q(x,v_k,t)| < \epsilon,$$

if $0 < |t| < \delta$. Note that the same $\delta > 0$ works for all $v \in \mathbb{S}^{m-1}$.

Now for any $y \in \mathbb{R}^m$ with $y \neq x$, we write $v = (y-x)/|y-x|$, so that $y = x + tv$ with $t = |y - x|$. Then

$$\begin{aligned}
&f(y) - f(x) - Df(x) \cdot (y-x) = f(x+tv) - f(x) - tv \cdot Df(x) \\
=&|y-x| Q(x, v, |y-x|) = o(|x-y|) \quad \text{as } y \to x.
\end{aligned}$$

Hence f is differentiable at x. □

Now, we state the following C^1 approximation result.

Theorem 4.6. *Suppose $f : \mathbb{R}^m \to \mathbb{R}$ is Lipschitz. Then for any $\epsilon > 0$, there exists a C^1 function $g : \mathbb{R}^m \to \mathbb{R}$ such that*

$$\mathcal{L}^m\big(\{x \in \mathbb{R}^m; f(x) \neq g(x)\} \cup \{x \in \mathbb{R}^m; Df(x) \neq Dg(x)\}\big) < \epsilon.$$

We skip the proof as it is quite complicated.

5. Area and Coarea Formulae

In this section, we discuss the area formula and coarea formula for Lipschitz functions on Euclidean spaces. We first need a result on polar decompositions from the linear algebra.

Lemma 5.1. *Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map.*

(i) If $m \leq n$, then there exist a symmetric map $S : \mathbb{R}^m \to \mathbb{R}^m$ and an orthogonal map $O : \mathbb{R}^m \to \mathbb{R}^n$ such that $L = O \circ S$.

(ii) If $m \geq n$, then there exist a symmetric map $S : \mathbb{R}^n \to \mathbb{R}^n$ and an orthogonal map $O : \mathbb{R}^m \to \mathbb{R}^n$ such that $L = S \circ O$.

For case (i), Theorem 5.1 asserts that a linear map can be decomposed as a composition of a dilation S and an orthogonal injection O. A similar interpretation is immediate for case (ii).

Proof. We need only prove (i). For (ii), we simply consider the adjoint map $L^* : \mathbb{R}^n \to \mathbb{R}^m$ and apply (i).

Let $A = L^* \circ L$. Then A is a symmetric nonnegative definite matrix. Let $\mu_{1,} \cdots , \mu_m \geq 0$ be the set of eigenvalues and $e_1, \cdots , e_m$ be a set of corresponding orthonormal eigenvectors, i.e., $Ae_k = \mu_k e_k$, $k = 1, \cdots , m$. Furthermore, let $\lambda_j = \sqrt{\mu_j}$ and $z_j = L(e_j/\lambda_j)$ for $\lambda_j \neq 0$. It is easy to see that $< z_k, z_l >= \delta_{kl}$ for any k and l with $\lambda_k \neq 0$ and $\lambda_l \neq 0$. Note that $\{z_j\}$ is defined only for j with $\lambda_j \neq 0$. Supplement those z_j by unit vectors such that $\{z_j\}_{j=1}^m$ forms an orthonormal set. Define $S : \mathbb{R}^m \to \mathbb{R}^m$ by $Se_j = \lambda_j e_j$ and define $O : \mathbb{R}^m \to \mathbb{R}^n$ by $Oe_j = z_j$, $j = 1, \cdots , m$. Then it is easy to check $L = O \circ S$. □

We define the *Jacobian* of $L : \mathbb{R}^m \to \mathbb{R}^n$ by

$$[[L]] = |\det S|.$$

It is easy to see that

$$[[L]] = \begin{cases} \sqrt{\det(L^*L)}, & m \leq n, \\ \sqrt{\det(LL^*)}, & n \leq m. \end{cases}$$

Hence $[[L]]$ is independent of the decomposition in Lemma 5.1. Obviously, $[[L]] = 0$ if $\text{rank}(L) < \min\{m, n\}$. By the Cauchy-Binet formula, we can calculate the Jacobian of $L : \mathbb{R}^m \to \mathbb{R}^n$, $m \leq n$, by

$$[[L]]^2 = \sum_{\lambda \in \Lambda(n,m)} \big(\det(P_\lambda \circ L)\big)^2,$$

where $\Lambda(n, m)$ is the (n, m)-permutation group and P_λ is the orthogonal projection to the appropriate m-dimensional subspace in $\mathbb{R}^n$.

Remark 5.2. Let $L_1, L_2 : \mathbb{R}^m \to \mathbb{R}^n$ be linear maps and $t > 0$. If $|L_1 v| \leq t|L_2 v|$ for any $v \in \mathbb{R}^m$, then $[[L_1]] \leq t^m [[L_2]]$. Such a simple fact will be employed frequently.

Now we introduce the Jacobian of differentiable maps. Suppose $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$ is a map differentiable at $x \in A$. We define the *Jacobian* of f at x by

$$Jf(x) = [[Df(x)]].$$

Sometimes, in order to emphasize the maximal possible rank, we denote the Jacobian by $J_m f(x)$ for $m \leq n$ and $J_n f(x)$ if $m \geq n$.

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a Lipschitz map and A be an $\mathcal{L}^m$-measurable subset of $\mathbb{R}^m$. We intend to calculate

$$\int_A Jf(x) d\mathcal{L}^m(x).$$

We will derive the area formula and the coarea formula for $m \leq n$ and $m \geq n$ respectively.

We first state state the area formula.

Theorem 5.3. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a Lipschitz map and $m \leq n$. Then for any $\mathcal{L}^m$-measurable set $A \subseteq \mathbb{R}^m$,*

$$\int_A J_m f(x) dx = \int_{\mathbb{R}^n} \mathcal{H}^0\big(A \cap f^{-1}(y)\big) d\mathcal{H}^m(y).$$

In particular, if f is one-to-one on A, then

$$\int_A J_m f(x) dx = \mathcal{H}^m\big(f(A)\big).$$

We note that the integral in the right hand side of the first formula is over $f(A)$. We will prove that it is an $\mathcal{H}^m$-measurable set of $\mathbb{R}^n$.

To prove Theorem 5.3, we need three preliminary lemmas.

Lemma 5.4. *Suppose $L : \mathbb{R}^m \to \mathbb{R}^n$ is linear, $m \leq n$. Then $\mathcal{H}^m(L(A)) = [[L]]\mathcal{L}^m(A)$.*

Proof. By Lemma 5.1, we have $L = O \circ S$, where $S : \mathbb{R}^m \to \mathbb{R}^m$ is symmetric and $O : \mathbb{R}^m \to \mathbb{R}^n$ is orthogonal. Then

$$\mathcal{H}^m\big(L(A)\big) = \mathcal{H}^m\big(O \circ S(A)\big) = \mathcal{H}^m\big(S(A)\big) = |\text{det} S|\mathcal{L}^m(A) = [[L]]\mathcal{L}^m(A).$$

□

Lemma 5.5. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz, $m \leq n$. If $A \subset \mathbb{R}^m$ is $\mathcal{L}^m$-measurable, then*

(i) $f(A)$ is an $\mathcal{H}^m$-measurable subset of $\mathbb{R}^n$;

(ii) the function $y \mapsto \mathcal{H}^0(A \cap f^{-1}(y))$ is $\mathcal{H}^m$-measurable on $\mathbb{R}^n$;

(iii) $\int_{\mathbb{R}^n} \mathcal{H}^0(f^{-1}(y) \cap A) d\mathcal{H}^m(y) \leq (Lipf)^m \mathcal{L}^m(A)$.

Proof. Without loss of generality, we assume that A is bounded. Otherwise, we simply decompose it into a countable union of bounded sets.

(i) For $i = 1, 2, \cdots$, we choose a sequence of compact sets $K_i \subset A$ such that $\mathcal{L}^m(K_i) > \mathcal{L}^m(A) - 1/i$. Since f is continuous, $f(K_i)$ is compact, which implies $f(K_i)$ is $\mathcal{H}^m$-measurable, and so is their countable union $\cup_{i=1}^\infty f(K_i) = f(\cup_{i=1}^\infty K_i)$. Then, we have

$$\mathcal{H}^m\big(f(A) \setminus f(\bigcup_{i=1}^\infty K_i)\big) \leq \mathcal{H}^m\big(f(A \setminus \bigcup_{i=1}^\infty K_i)\big) \leq (\text{Lip} f)^m \mathcal{L}^m(A \setminus \bigcup_{i=1}^\infty K_i) = 0.$$

This implies that $f(A)$ is $\mathcal{H}^m$-measurable.

(ii) We first decompose $\mathbb{R}^m$ into a set of pairwise disjoint small cubes. For any integer $k \geq 1$, let $\mathcal{B}_k = \{Q; Q = (a_1, b_1] \times \cdots \times (a_m, b_m]\}$, where $a_i = n_i/2^k$, $b_i = (n_i+1)/2^k$, n_i's are integers. Note that any element of $\mathcal{B}_{k+1}$ is a subset of some element of $\mathcal{B}_k$. Obviously, $\mathbb{R}^m = \cup_{Q\in\mathcal{B}_k} Q$. Define

$$g_k = \sum_{Q\in\mathcal{B}_k} \chi_{f(A\cap Q)} \quad \text{in } \mathbb{R}^n.$$

Then by (i), g_k is an $\mathcal{H}^m$-measurable function in $\mathbb{R}^n$ for any k. Note that $g_k(y)$ is the number of cubes $Q \in \mathcal{B}_k$ such that $f^{-1}(y) \cap A \cap Q$ is nonempty. Therefore, g_k increases monotonically to $\mathcal{H}^0(f^{-1}(y) \cap A)$ as $k \to \infty$. By the Monotone Convergence Theorem, $\mathcal{H}^0(f^{-1}(y) \cap A)$ is also an $\mathcal{H}^m$-measurable function on $\mathbb{R}^n$.

(iii) We first note that the result of (iii) is sharper than Lemma 4.2 for $s = m$. Let g_k be the functions defined in (ii). Then, by the Monotone Convergence Theorem, we have

$$\begin{aligned}
&\int_{\mathbb{R}^n} \mathcal{H}^0(f^{-1}(y) \cap A) d\mathcal{H}^m(y) = \lim_{k\to\infty} \int_{\mathbb{R}^n} g_k(y) d\mathcal{H}^m(y) \\
=&\lim_{k\to\infty} \sum_{Q\in\mathcal{B}_k} \mathcal{H}^m\big(f(A \cap Q)\big) \leq \lim_{k\to\infty} \sum_{Q\in\mathcal{B}_k} (\text{Lip} f)^m \mathcal{L}^m(Q \cap A) \\
=&(\text{Lip} f)^m \mathcal{L}^m(A).
\end{aligned}$$

This finishes the proof. □

Lemma 5.6. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz, $m \leq n$, $t > 1$ and*

$$\mathcal{R} = \{x; Df(x) \text{ exists and } J_m f(x) > 0\}.$$

Then there is a countable collection $\{E_k\}_{k=1}^\infty$ of Borel sets of $\mathbb{R}^m$ such that

(i) $\mathcal{R} = \cup_{k=1}^\infty E_k$;

(ii) $f |_{E_k}$ is one-to-one;

(iii) for each $k = 1, 2, \cdots$, there is a symmetric nonsingular linear map $T_k : \mathbb{R}^m \to \mathbb{R}^m$ such that

$$Lip(f |_{E_k} \circ T_k^{-1}) \leq t, \quad Lip(T_k \circ (f |_{E_k})^{-1}) \leq t,$$

and

$$\begin{aligned}
t^{-1}|T_k v| \leq &|Df(p) \cdot v| \leq t|T_k v| \quad \text{for any } p \in E_k,\ v \in \mathbb{R}^m, \\
t^{-m}|detT_k| &\leq (J_m f) |_{E_k} \leq t^m |detT_k|.
\end{aligned}$$

Lemma 5.6 asserts that f can be locally approximated by a nonsingular symmetric linear map as close as we wish.

Proof. We fix a small $\epsilon > 0$ so that ϵ satisfies $\frac{1}{t} + \epsilon < 1 < t - \epsilon$. Let $\mathcal{C}$ be a countable dense subset of $\mathcal{R}$ and let $\mathcal{S}$ be a countable dense subset of symmetric nonsingular linear maps of $\mathbb{R}^m$. Then we define, for each $c \in \mathcal{C}$, $T \in \mathcal{S}$, $i = 1, 2, \cdots$, a set $E(c, T, i)$ to be all $b \in \mathcal{R} \cap B_{1/i}(c)$ satisfying

$$(\frac{1}{t} + \epsilon)|Tv| \leq |Df(b) \cdot v| \leq (t - \epsilon)|Tv| \quad \text{for any } v \in \mathbb{R}^m, \tag{5.1}$$

and

$$|f(a) - f(b) - Df(b) \cdot (a - b)| \leq \epsilon |T(a - b)| \quad \text{for any } a \in B_{\frac{2}{i}}(b). \tag{5.2}$$

Note that for any $b \in \mathcal{R}$, we have $Df(b) = O \circ S$, for an orthogonal map $O : \mathbb{R}^m \to \mathbb{R}^n$ and a nonsingular symmetric map S on $\mathbb{R}^m$, and hence

$$|f(a) - f(b) - Df(b) \cdot (a - b)| \leq \epsilon |S(a - b)|,$$

if $|a - b|$ is small. Since $\mathcal{S}$ is dense in the set of all symmetric nonsingular maps, we can replace S by a $T \in \mathcal{S}$ so that (5.1) and (5.2) are satisfied. Thus $b \in E(c, T, i)$ for some c, T, i.

Also note that $E(c, T, i)$ is a Borel set since Df is Borel measurable. By (5.1) and (5.2), we have

$$\frac{1}{t}|T(a - b)| \leq |f(a) - f(b)| \leq t|T(a - b)|,$$

for any $b \in E(c, T, i)$ and $a \in B_{2/i}(b)$. This shows in particular that f is one-to-one on $E(c, T, i)$. By (5.1), we have

$$\frac{1}{t}|Tv| \leq |Df(b) \cdot v| \leq t|Tv| \quad \text{for any } v \in \mathbb{R}^m,$$

and hence

$$t^{-m}|\text{det}T| \leq J_m f(b) \leq t^m |\text{det}T|,$$

for any $b \in E(c, T, i)$. We relabel $E(c, T, i)$ as $\{E_k\}_{k=1}^{\infty}$ and this concludes the proof. □

Now we are ready to prove Theorem 5.3.

Proof of Theorem 5.3. Since Df exists almost everywhere, by Lemma 5.5 (iii), we assume that Df exists for all $x \in A$.

Case 1. Suppose $A \subset \{J_m f(x) > 0\}$.

Fix a $t > 1$ and choose Borel sets $\{E_i\}_{i=1}^{\infty}$ as in Lemma 5.6. We assume that all of E_i are pairwise disjoint. For each fixed integer $k \geq 1$, we define $\mathcal{B}_k$ as in Lemma 5.5 and set $F_j^i = E_j \cap Q_i \cap A$, $Q_i \in \mathcal{B}_k$. Then $\{F_j^i\}$ is disjoint and $A = \cup_{i,j} F_j^i$.

We claim that

$$\lim_{k \to \infty} \sum_{i,j=1}^{\infty} \mathcal{H}^m\big(f(F_j^i)\big) = \int_{\mathbb{R}^n} \mathcal{H}^0\big(f^{-1}(y) \cap A\big) d\mathcal{H}^m(y). \tag{5.3}$$

To prove this, let $g_k = \sum_{i,j} \chi_{f(F_j^i)}$ in $\mathbb{R}^n$. Then g_k is $\mathcal{H}^m$-measurable by Lemma 5.5 and $g_k(y)$ is the number of sets $\{F_j^i\}$ such that $F_j^i \cap f^{-1}(y) \neq \emptyset$ for any $y \in \mathbb{R}^n$. Thus $g_k(y)$ converges to $\mathcal{H}^0(f^{-1}(y) \cap A)$ monotonically. By applying the Monotone Convergence Theorem, we obtain (5.3).

Next, we have by Lemma 5.6

$$\begin{aligned}\mathcal{H}^m\big(f(F_j^i)\big) &= \mathcal{H}^m\big(f\mid_{E_j}\circ T_j^{-1}\circ T_j(F_j^i)\big)\\ &\le \big(\mathrm{Lip}(f\mid_{E_j}\circ T_j^{-1})\big)^m\mathcal{H}^m\big(T_j(F_j^i)\big)\\ &\le t^m\mathcal{H}^m\big(T_j(F_j^i)\big) = t^m|\det T_j|\mathcal{L}^m(F_j^i)\\ &\le t^{2m}\int_{F_j^i} J_mf(x)dx\\ &\le t^{3m}\int_{F_j^i}|\det T_j|dx \le t^{3m}|\det T_j|\mathcal{L}^m(F_j^i)\\ &\le t^{4m}\mathcal{H}^m\big(f(F_j^i)\big).\end{aligned}$$

Thus by dividing by t^{2m} and summing over i and j, we obtain

$$t^{-2m}\sum_{i,j}\mathcal{H}^m\big(f(F_j^i)\big) \le \int_A J_mf(x)dx \le t^{2m}\sum_{i,j}\mathcal{H}^m\big(f(F_j^i)\big).$$

By taking $t \to 1$, we conclude the Case 1.

Case 2. Suppose $A \subset \{J_mf(x) = 0\}$.

Fix an $\epsilon > 0$. Define $g : \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ by $g(x) = (f(x), \epsilon x)$ and $P : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ by the projection $P(y,z) = y$ for any $(y,z) \in \mathbb{R}^n \times \mathbb{R}^m$. We then have $f(x) = P \circ g(x)$. We can check easily that $0 < J_mg(x) < C\epsilon$ on A, for a constant C. Since $\mathrm{Lip}P = 1$, we have by Case 1,

$$\begin{aligned}\mathcal{H}^m(f(A)) \le \mathcal{H}^m(g(A)) &\le \int_{\mathbb{R}^{m+n}}\mathcal{H}^0\big(A\cap g^{-1}(y,z)\big)d\mathcal{H}^m(y,z)\\ &= \int_A J_mg(x)dx \le C\epsilon\mathcal{L}^m(A) \to 0 \quad \text{as } \epsilon \to 0.\end{aligned}$$

Thus we obtain by letting $\epsilon \to 0$,

$$\mathcal{H}^m\big(f(A)\big) = 0,$$

and hence

$$\int_{\mathbb{R}^n}\mathcal{H}^0\big(A\cap f^{-1}(y)\big)d\mathcal{H}^m(y) = \int_{f(A)}\mathcal{H}^0\big(A\cap f^{-1}(y)\big)d\mathcal{H}^m(y) = 0.$$

Theorem 5.3 now is obtained by combining Cases 1 and 2. □

Corollary 5.7. *Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz and $m \le n$. For every $\mathcal{L}^m$-measurable set A, there exists a Borel set*

$$B \subset \{x \in A; J_mf(x) > 0\}$$

such that $f|_B$ is one-to-one and $\mathcal{H}^m\big(f(A)\setminus f(B)\big) = 0$.

Proof. In case $\mathcal{L}^m(A) = 0$, we take $B = \emptyset$.

In case A is a Borel set, we use Lemma 5.6 to obtain Borel sets E_i such that $f|_{E_i}$ is one-to-one for $i = 1, 2, \cdots$, and

$$P = \{x \in A; J_mf(x) > 0\} \subset \cup_{i=1}^\infty E_i.$$

By setting

$$F_i = P \cap E_i \setminus \bigcup_{j=1}^{i-1} f^{-1}\big(f(P \cap E_j)\big) \quad \text{for } i = 1, 2, \cdots,$$
$$B = \cup_{i=1}^{\infty} F_i,$$

we conclude by Theorem 5.3 easily that B is Borel, $f|_B$ is one-to-one, $f(B) = f(P)$ and

$$\mathcal{H}^m\big(f(A) \setminus f(B)\big) \le \mathcal{H}^m\big(f(A \setminus P)\big) \le \int_{A \setminus P} J_m f d\mathcal{L}^m = 0.$$

This finishes the proof. □

We can also obtain the following change of variable formula.

Theorem 5.8. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a Lipschitz function, $m \le n$, and let $u \in L^1(\mathbb{R}^m)$. Then*

$$\int_{\mathbb{R}^m} u(x) J_m f(x) dx = \int_{\mathbb{R}^n} \sum_{x \in f^{-1}(y)} u(x) d\mathcal{H}^m(y). \tag{5.4}$$

In particular, if f is one-to-one, then

$$\int_{\mathbb{R}^m} u(x) J_m f(x) dx = \int_{\mathbb{R}^n} u\big(f^{-1}(y)\big) d\mathcal{H}^m(y).$$

Proof. We consider the case $u \ge 0$. The general case then follows easily.

We first consider a decomposition of u. Define

$$A_1 = \{u \ge 1\},$$

and inductively

$$A_k = \Big\{u \ge \frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j} \chi_{A_j}\Big\} \quad \text{for } k > 1.$$

We claim

$$u = \sum_{i=1}^{\infty} \frac{1}{i} \chi_{A_i}.$$

In fact, it is easy to see that (i) $u \ge \sum_{j=1}^{k} \frac{1}{j} \chi_{A_j}$ and (ii), for $0 < u(x) < \infty$ and for sufficiently large k, $x \notin A_k$ implies $0 \le u(x) - \sum_{j=1}^{k-1} \frac{1}{j} \chi_{A_j}(x) \le \frac{1}{k}$. Now the above claim follows from (i) and (ii) readily. With this decomposition, we have

$$\int_{\mathbb{R}^m} u(x) J_m f(x) dx = \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^m} \chi_{A_i} J_m f(x) dx$$
$$= \sum_{i=1}^{\infty} \frac{1}{i} \int_{A_i} J_m f(x) dx = \sum_{i=1}^{\infty} \frac{1}{i} \int_{\mathbb{R}^n} \mathcal{H}^0(f^{-1}(y) \cap A_i) d\mathcal{H}^m(y),$$

where we used Theorem 5.3 in the last equality. This implies

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathcal{H}^0\big(f^{-1}(y) \cap A_i\big)$$

is finite for $\mathcal{H}^m$-almost all $y \in \mathbb{R}^n$. By the claim, it is equal to

$$\sum_{x \in f^{-1}(y)} u(x).$$

This finishes the proof. □

In the rest of this section section, we discuss the coarea formula for Lipschitz functions on Euclidean spaces. We first state the result.

Theorem 5.9. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz and $m \geq n$. Then for any $\mathcal{L}^m$-measurable subset $A \subset \mathbb{R}^m$,*

$$\int_A J_n f(x) dx = \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) dy.$$

The co-area formula is the generalized version of Fubini Theorem. This can be seen easily by choosing f to be an orthogonal projection.

Remark 5.10. If we take $A = \{x \in \mathbb{R}^m; J_n f(x) = 0\}$, the set of critical points, we obtain for $\mathcal{L}^n$-almost all $y \in \mathbb{R}^n$,

$$\mathcal{H}^{m-n}\big(\{x \in \mathbb{R}^m; J_n f(x) = 0\} \cap f^{-1}(y)\big) = 0.$$

This statement is weaker than the standard Sard Theorem.

Similar to Theorem 5.8, we also have the following change of variable formula for the coarea formula.

Theorem 5.11. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a Lipschitz function, $m \geq n$, and let $u \in L^1(\mathbb{R}^m)$. Then*

$$\int_{\mathbb{R}^m} u(x) J_n f(x) dx = \int_{\mathbb{R}^n} \int_{f^{-1}(y)} u d\mathcal{H}^{m-n} dy.$$

Now we begin to prove Theorem 5.9. First, we need a few lemmas.

Lemma 5.12. *Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map, $m \geq n$, and $A \subset \mathbb{R}^m$ be an $\mathcal{L}^m$-measurable subset. Then*

(i) the map $y \mapsto \mathcal{H}^{m-n}(A \cap L^{-1}(y))$ is $\mathcal{L}^n$-measurable;

(ii) $\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap L^{-1}(y)) dy = [[L]] \mathcal{L}^m(A)$.

Proof. We divide the proof into the following three cases.

Case 1. Suppose $\dim L(\mathbb{R}^m) < n$, i.e., $[[L]] = 0$. Then $A \cap L^{-1}(y) = \emptyset$ and hence $\mathcal{H}^{m-n}(A \cap L^{-1}(y)) = 0$ for $\mathcal{L}^n$-almost all $y \in \mathbb{R}^n$. Thus (i) and (ii) are trivial.

Case 2. Suppose L is an orthogonal projection $P : \mathbb{R}^m \to \mathbb{R}^n$. Then, by Fubini Theorem, $y \mapsto \mathcal{L}^{m-n}(A \cap P^{-1}(y))$ is $\mathcal{L}^n$-measurable and $\int_{\mathbb{R}^n} \mathcal{L}^{m-n}(A \cap P^{-1}(y)) dy = \mathcal{L}^m(A)$. So (i) and (ii) are proved for this case.

Case 3. Suppose $L = S \circ O$ where $O : \mathbb{R}^m \to \mathbb{R}^n$ is an orthogonal projection and $S : \mathbb{R}^n \to \mathbb{R}^n$ is a symmetric nonsingular map as in Lemma 5.1. Then by Case 2 and by the change of variables, we have

$$\begin{aligned}\mathcal{L}^m(A) &= \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap O^{-1}(y)\big)dy \\ &= \det S^{-1} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap O^{-1} \circ S^{-1}(z)\big)dz \quad (\text{where } z = Sy) \\ &= \det S^{-1} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap L^{-1}(z))dz.\end{aligned}$$

This concludes the proof. □

Lemma 5.13. *Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be Lipschitz for $m \geq n$ and $A \subset \mathbb{R}^m$ be an $\mathcal{L}^m$-measurable set. Then*

(i) $f(A)$ is an $\mathcal{L}^n$-measurable subset of $\mathbb{R}^n$;

(ii) the function $y \mapsto \mathcal{H}^{m-n}(A \cap f^{-1}(y))$ is $\mathcal{L}^n$-measurable on $\mathbb{R}^n$;

(iii) $\displaystyle\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y))dy \leq \frac{\omega(m-n)\omega(n)}{\omega(m)}(Lipf)^n \mathcal{L}^m(A)$.

Proof. The proof of (i) is the same as in the corresponding case for the area formula in Lemma 5.5.

Next, we prove that A satisfies (iii) if it satisfies (ii). For each $j = 1, 2, \cdots$, there is a collection of closed balls $\{B_i^j\}_{i=1}^\infty$ such that $A \subset \cup_{i=1}^\infty B_i^j$, $\mathrm{diam} B_i^j \leq 1/j$ and $\sum_{i=1}^\infty \mathcal{L}^m(B_i^j) \leq \mathcal{L}^m(A) + 1/j$. Define

$$g_i^j(y) = \omega(m-n)\big(\frac{1}{2}\mathrm{diam} B_i^j\big)^{m-n} \chi_{f(B_i^j)}(y).$$

By (i), g_i^j is an $\mathcal{L}^n$-measurable function on $\mathbb{R}^n$, and for any $y \in \mathbb{R}^n$

$$\mathcal{H}^{m-n}_{\frac{1}{j}}\big(A \cap f^{-1}(y)\big) \leq \sum_{i=1}^\infty g_i^j(y).$$

By Fatou's Lemma and (ii), which we assumed to be true, we obtain

$$\begin{aligned}&\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big)dy = \int_{\mathbb{R}^n} \lim_{j\to\infty} \mathcal{H}^{m-n}_{\frac{1}{j}}\big(A \cap f^{-1}(y)\big)dy \\ &\leq \int_{\mathbb{R}^n} \liminf_{j\to\infty} \sum_{i=1}^\infty g_i^j(y)dy \leq \liminf_{j\to\infty} \sum_{i=1}^\infty \int_{\mathbb{R}^n} g_i^j(y)dy \\ &= \liminf_{j\to\infty} \sum_{i=1}^\infty \omega(m-n)\big(\frac{1}{2}\mathrm{diam} B_i^j\big)^{m-n} \mathcal{L}^n\big(f(B_i^j)\big) \\ &\leq \liminf_{j\to\infty} \sum_{i=1}^\infty \omega(m-n)\big(\frac{1}{2}\mathrm{diam} B_i^j\big)^{m-n} (\mathrm{Lip} f)^n \omega(n)\big(\frac{1}{2}\mathrm{diam} B_i^j\big)^n \\ &\leq \frac{\omega(n)\omega(m-n)}{\omega(m)}(\mathrm{Lip} f)^n \mathcal{L}^m(A),\end{aligned}$$

where we used the isodiametric inequality

$$\mathcal{L}^n\big(f(B_i^j)\big) \le \omega(n)\big(\frac{1}{2}\mathrm{diam} f(B_i^j)\big)^n.$$

To prove (ii), we first consider the case when A is compact. For each fix $t \ge 0$, we claim

$$\{y \in \mathbb{R}^n; \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) \le t\} = \bigcap_{i=1}^{\infty} U_i(t),$$

where, for each integer $i \ge 1$, $U_i(t)$ is the collection of those points $y \in \mathbb{R}^n$ such that there are finitely many open sets $S_1, \cdots, S_l$, $A \cap f^{-1}(y) \subset \cup_{j=1}^{l} S_j$, diam $S_j < 1/i, j = 1, \cdots, l$, and $\displaystyle\sum_{j=1}^{l} \omega(m-n) \cdot (\mathrm{diam} S_j/2)^{m-n} \le t + \frac{1}{i}$. Since f is continuous, $U_i(t)$ is open. This shows that $y \mapsto \mathcal{H}^{m-n}(A \cap f^{-1}(y))$ is Borel measurable. Now we prove the claim. If $\mathcal{H}^{m-n}(A \cap f^{-1}(y)) \le t$, then $\mathcal{H}_\delta^{m-n}(A \cap f^{-1}(y)) \le t$ for any $\delta > 0$. Given any integer $i \ge 1$, there exists a collection of sets $\{S_j\}_{j=1}^{\infty}$ such that $A \cap f^{-1}(y) \subset \cup_{j=1}^{\infty} S_j$, diam$S_j < 1/i$ and $\sum_{j=1}^{\infty} \omega(m-n)(\mathrm{diam} S_j/2)^{m-n} < t + 1/i$. Furthermore, we can assume that S_j are open. Since $A \cap f^{-1}(y)$ is compact, we can choose a finite number of S_j. This shows that $y \in U_i(t)$, i.e.,

$$\{y \in \mathbb{R}^n; \mathcal{H}^{m-n}(A \cap f^{-1}(y)) \le t\} \subset \bigcap_{i=1}^{\infty} U_i(t).$$

The other inclusion is immediate, and this proves the claim.

Next, we treat the case when A is open. A can be written as a countable union of compact sets, i.e., $A = \cup_{i=1}^{\infty} K_i$ with compact sets $K_i \subset K_{i+1}$, $i = 1, 2, \cdots$. Then

$$\mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) = \lim_{i \to \infty} \mathcal{H}^{m-n}\big(K_i \cap f^{-1}(y)\big).$$

Thus $\mathcal{H}^{m-n}(A \cap f^{-1}(y))$ is Borel measurable on $\mathbb{R}^n$.

To show (ii) for the general case, let K_i and U_i be compact and open sets such that $K_i \subset A \subset U_i$ and $\mathcal{L}^m(U_i \backslash K_i) < 1/i$. Note (iii) holds for $U_i \setminus K_i$. Then we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n} \big(\mathcal{H}^{m-n}(U_i \cap f^{-1}(y)) - \mathcal{H}^{m-n}(K_i \cap f^{-1}(y))\big) dy \\
=&\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big((U_i \backslash K_i) \cap f^{-1}(y)\big) dy \\
\le&\frac{\omega(n)\omega(m-n)}{\omega(m)} (\mathrm{Lip} f)^n \mathcal{L}^m(U_i \backslash K_i) \to 0,
\end{aligned}$$

as $i \to \infty$. Since

$$\mathcal{H}^{m-n}\big(K_i \cap f^{-1}(y)\big) \le \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) \le \mathcal{H}^{m-n}\big(U_i \cap f^{-1}(y)\big),$$

we see that $\mathcal{H}^{m-n}(A \cap f^{-1}(y))$ is a measurable function of y. □

Now we begin to prove Theorem 5.9.

Proof of Theorem 5.9. First assume that $\mathcal{L}^m(A) < \infty$ and $Df(x)$ exists for each $x \in A$, since measure zero sets can be ignored by Lemma 5.13(iii).

Case 1. Suppose $A \subset \{x \in \mathbb{R}^m; J_n f(x) > 0\}$.

For a $\lambda \in \Lambda(m, m-n)$, let $P_\lambda : \mathbb{R}^m \to \mathbb{R}^{m-n}$ be the orthogonal projection defined by $P(x) = (x_{\lambda_1}, \cdots, x_{\lambda_{m-n}})$. Then $f(x) = q \circ h_\lambda(x)$, where $h_\lambda : \mathbb{R}^m \to \mathbb{R}^m$ is defined by $h_\lambda(x) = (f(x), P_\lambda(x))$ and $q : \mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n} \to \mathbb{R}^n$ is defined by $q(y, z) = y$.

Let $A_\lambda = \{x \in A; J_m h_\lambda(x) > 0\}$. Then $A = \cup_{\lambda \in \Lambda(m,m-n)} A_\lambda$. So we assume that $A = A_\lambda$ for a $\lambda \in \Lambda(m, m-n)$ and denote h_λ by h. Moreover, we assume that h is one-to-one and $|h(x) - h(y)| \geq c|x - y|$ for any $x, y \in A$. In fact, by Lemma 5.6, A may be decomposed into a disjoint union of Borel sets $\{F_l\}_{l=1}^\infty$ such that h is one-to-one in each F_l and $|h(x) - h(y)| \geq c_l|x - y|$ for any $x, y \in F_l$. So h^{-1} is Lipschitz in $h(F_l)$.

We apply Lemma 5.6 to h^{-1} in $h(A)$. For each fixed $t > 1$, there exists a disjoint union $A = \cup_{k=1}^\infty E_k$ and associated symmetric nonsingular linear maps $S_k : \mathbb{R}^m \to \mathbb{R}^m$, $k = 1, 2, \cdots$, such that

$$\mathrm{Lip}(S_k^{-1} \circ h \,|_{E_k}) \leq t, \quad \mathrm{Lip}((h\,|_{E_k})^{-1} \circ S_k) \leq t,$$

and

$$t^{-1}|S_k^{-1}v| \leq |Dh^{-1}(y) \cdot v| \leq t|S_k^{-1}v| \quad \text{for any } y \in h(E_k) \text{ and } v \in \mathbb{R}^m.$$

Then we have

$$\begin{aligned}
&\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(E_k \cap f^{-1}(y)\big)dy = \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(E_k \cap h^{-1} \circ q^{-1}(y)\big)dy \\
=&\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(h^{-1} \circ S_k \circ S_k^{-1}(h(E_k) \cap q^{-1}(y))\big)dy \\
\leq& t^{m-n} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(S_k^{-1} \circ h(E_k) \cap S_k^{-1} \circ q^{-1}(y)\big)dy \\
=& t^{m-n}[[q \circ S_k]]\mathcal{L}^m\big(S_k^{-1} \circ h(E_k)\big) \quad \text{(by Lemma 5.12(ii))} \\
\leq& t^{2m-n}[[q \circ S_k]]\mathcal{L}^m(E_k) \\
\leq& t^{2m} \int_{E_k} J_m f(x)dx,
\end{aligned}$$

where in the last step we used $[[q \circ S_k]] \leq t^n J_n f(x)$ for any $x \in E_k$. Similarly, we have

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(E_k \cap f^{-1}(y)\big)dy \geq t^{-2m} \int_{E_k} J_n f(x)dx.$$

To conclude Case 1, we sum over k and let $t \to 1$.

Case 2. Suppose $A \subset \{x \in \mathbb{R}^m; J_n f(x) = 0\}$. We prove $\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(f^{-1}(y) \cap A)dy = 0$.

Fix an $\epsilon > 0$ and define

$$\begin{aligned}
g &: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n \ \text{ by } g(x, z) = f(x) + \epsilon z, \\
P &: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n \ \text{ by } P(x, z) = z.
\end{aligned}$$

Then we see $0 < J_n g(x,z) < C\epsilon$. Apply Case 1 to g in $A \times B_1^n \subset \mathbb{R}^m \times \mathbb{R}^n$, where B_1^n is the unit ball in $\mathbb{R}^n$. We get

$$\begin{aligned} C\epsilon \cdot \mathcal{L}^m(A)\mathcal{L}^n(B_1^n) \geq & \int_{A\times B_1^n} J_n g(x,z) d\mathcal{L}^{m+n}(x,z) \\ = & \int_{\mathbb{R}^n} \mathcal{H}^m\big((A \times B_1^n) \cap g^{-1}(y)\big) d\mathcal{L}^n(y). \end{aligned}$$

A slight modification of Lemma 5.13 shows that

$$\begin{aligned} &\mathcal{H}^m\big((A \times B_1^n) \cap g^{-1}(y)\big) \\ \geq & c(m,n) \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big((A \times B_1^n) \cap g^{-1}(y) \cap P^{-1}(w)\big) d\mathcal{L}^n(w). \end{aligned}$$

Thus

$$\begin{aligned} &c\epsilon\mathcal{L}^m(A)\mathcal{L}^n(B_1^n) \\ \geq & c \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big((A \times B_1^n) \cap g^{-1}(y) \cap P^{-1}(w)\big) d\mathcal{L}^n(\omega) d\mathcal{L}^n(y), \end{aligned} \tag{5.5}$$

where

$$(x,z) \in A \times B_1^n,\ g(x,z) = y,\ p(x,z) = w,$$

which is equivalent to $x \in f^{-1}(y - \epsilon w)$ and $w \in B_1^n$. Hence,

$$\begin{aligned} &\text{the right side of (5.5)} \\ = & c \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y - \epsilon w)\big) d\mathcal{L}^n(w)\mathcal{L}^n(y) \\ = & c \int_{B_1^n}\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y - \epsilon w)\big) d\mathcal{L}^n(y) d\mathcal{L}^n(w) \\ = & c\omega(n) \int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) d\mathcal{L}^n(y). \end{aligned}$$

Letting $\epsilon \to 0$, we obtain

$$\int_{\mathbb{R}^n} \mathcal{H}^{m-n}\big(A \cap f^{-1}(y)\big) d\mathcal{L}^n(y) = 0.$$

Combining cases 1 and 2, we finish the proof. □

6. Countably Rectifiable Sets

Countably rectifiable sets provide an appropriate notion of *generalized submanifolds in Euclidean spaces*; they are the sets on which rectifiable currents and varifolds live. Throughout this section, m and n are two positive integers with $1 \leq m \leq n-1$.

Definition 6.1. An $\mathcal{H}^m$-measurable set $E \subset \mathbb{R}^n$ is countably m-rectifiable if

$$E \subset E_0 \cup \big(\bigcup_{j=1}^{\infty} f_j(\mathbb{R}^m)\big),$$

where $\mathcal{H}^m(E_0) = 0$ and $f_j : \mathbb{R}^m \to \mathbb{R}^n$ is a Lipschitz function for $j = 1, 2, \cdots$.

By Theorem 4.3, this is equivalent to

$$E = E_0 \cup \big(\bigcup_{j=1}^{\infty} f_j(A_j) \big),$$

where $\mathcal{H}^m(E_0) = 0$ and $f_j : A_j \to \mathbb{R}^n$ is a Lipschitz function on an $\mathcal{L}^m$-measurable subset $A_j \subset \mathbb{R}^m$, $j = 1, 2, \cdots$.

Definition 6.2. An $\mathcal{H}^m$-measurable set $E \subset \mathbb{R}^n$ is purely m-unrectifiable if E contains no countably m-rectifiable subsets of positive $\mathcal{H}^m$-measure.

Remark 6.3. It is useful to remark that there is a dichotomy between rectifiable and purely unrectifiable sets. Precisely speaking, for any $\mathcal{H}^m$-measurable subset $A \subset \mathbb{R}^n$, by using the Hausdorff maximal principle, there holds

$$A = B \cup C, \quad B \cap C = \emptyset,$$

where B is countably m-rectifiable and C is purely m-unrectifiable in $\mathbb{R}^n$.

Now, we introduce tangent cones. For any $p \in \mathbb{R}^n$, $v \in \mathbb{R}^n$ and $\varepsilon > 0$, we define

$$X(p, v, \varepsilon) = \{x \in \mathbb{R}^n; |r(x - p) - v| < \varepsilon \text{ for some } r > 0\}.$$

Note that the requirement on x can be written as $r(x - p) - v = \varepsilon y$ for some $y \in B_1$, or

$$x = p + \frac{v + \varepsilon y}{r}.$$

It is easy to see that $X(p, v, \varepsilon)$ is a cone and $X(p, sv, s\varepsilon) = X(p, v, \varepsilon)$ for any $s > 0$.

Definition 6.4. Suppose E is a subset of $\mathbb{R}^n$ and p is an arbitrary point in $\mathbb{R}^n$. (Note that p is not necessarily on E.) The *tangent cone* T_pE of E at p is defined by

$$T_pE = \{v \in \mathbb{R}^n; E \cap X(p, v, \varepsilon) \cap B_\varepsilon(p) \setminus \{p\} \neq \emptyset \text{ for any } \varepsilon > 0\}.$$

Vectors in T_pE are called *tangent vectors* of E at p.

Usually, we call T_xE the tangent space if it is a subspace of $\mathbb{R}^n$.

Alternatively, tangent vectors can be defined in the following way: $v \in T_pE \setminus \{0\}$ if and only if there exist $x_i \in E \setminus \{p\}$ and $r_i > 0$ such that $x_i \to p$ and $r_i(x - p) \to v$. If, in addition, $|v| = 1$, we may take $r_i = \dfrac{1}{|x_i - p|}$.

Remark 6.5. We observe that

(1) T_pE is a closed subset of $\mathbb{R}^n$;
(2) T_pE is a cone in $\mathbb{R}^n$, i.e., $v \in T_pE$ and $s \in [0, \infty)$ imply $sv \in T_pE$;
(3) $0 \in T_pE$ (or $T_pE \neq \emptyset$) if and only if $p \in \bar{E}$.
(4) $T_pE = T_p\bar{E}$.

Lemma 6.6. *Suppose E is a subset of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^l$ is a map differentiable at $p \in \mathbb{R}^n$. Then*

$$Df(p)\big(T_pE\big) \subset T_{f(p)}f(E).$$

Moreover, equality holds if $f|_E$ is one-to-one, $p \in E$, $(f|_E)^{-1}$ is continuous at $f(p)$, and $Df(p)$ is a linear homeomorphism mapping $\mathbb{R}^n$ onto a subspace of $\mathbb{R}^l$.

Proof. We assume $p = 0$, $f(p) = 0$ and write $Df(p) = L$. For each $v \in T_0E$, there exist $x_i \in E$, $r_i \in (0,\infty)$ with $x_i \to 0$, $r_ix_i \to v$ as $i \to \infty$. Hence

$$Lv - r_if(x_i) = L(v - r_ix_i) + |r_ix_i| \cdot \frac{Lx_i - f(x_i)}{|x_i|} \to 0 \quad \text{as } i \to \infty,$$

and $Lv \in T_0f(E)$. On the other hand, in case the additional hypotheses hold and $w \in T_0f(E)$, there exist $x_i \in E$, $r_i \in (0,\infty)$ with $f(x_i) \to 0$, $r_if(x_i) \to w$ as $i \to \infty$. Hence $x_i \to 0$ and

$$L(r_ix_i) - w = |r_if(x_i)| \cdot \frac{|x_i|}{|f(x_i)|} \cdot \frac{Lx_i - f(x_i)}{|x_i|} + r_if(x_i) - w \to 0 \text{ as } i \to \infty.$$

This implies that r_ix_i converges to some $v \in \mathbb{R}^n$ with $Lv = w$, and hence $v \in T_0E$. □

Of particular interest is the tangent space of an m-dimensional submanifold of $\mathbb{R}^n$.

Proposition 6.7. *Let $M \subset \mathbb{R}^n$ be a subset. Then the following statements are equivalent.*

(1) For any $p \in M$, there are neighborhoods U of p and V of 0 in $\mathbb{R}^n$ and a C^k diffeomorphism $f : U \to V$ such that

$$f(U \cap M) = V \cap (\mathbb{R}^m \times \{0\}).$$

(2) For any $p \in M$, there is a neighborhood U of p in $\mathbb{R}^n$ and a C^k map $h : U \to \mathbb{R}^{n-m}$ such that $U \cap M = h^{-1}(0)$ and h is a submersion at each point of U, i.e., Dh is surjective at each point of U.

(3) For any $p \in M$, there is a neighborhood U of p in $\mathbb{R}^n$, a neighborhood Ω of 0 in $\mathbb{R}^m$ and a C^k map $g : \Omega \to \mathbb{R}^n$ such that (Ω, g) is a local parametrization of $U \cap M$ around p, i.e., $g : \Omega \to M \cap U$ is a homeomorphism and g is an immersion at each point of Ω, i.e., Dg is injective at each point of Ω.

The subset M in Proposition 6.7 is called an *m-dimensional C^k submanifold* in $\mathbb{R}^n$.

Proposition 6.8. *Let $M \subset \mathbb{R}^n$ be an m-dimensional C^1 submanifold and $p \in M$. Suppose U is a neighborhood of p in $\mathbb{R}^n$.*

(1) Suppose f is a diffeomorphism from U onto a neighborhood V of 0 in $\mathbb{R}^n$ such that $f(U \cap M) = V \cap (\mathbb{R}^m \times \{0\})$. Then $T_pM = \big(Df(p)\big)^{-1}(\mathbb{R}^m \times \{0\})$.

(2) Suppose h is a submersion of U to $\mathbb{R}^{n-m}$ such that $U \cap M = h^{-1}(0)$. Then $T_pM = Ker\big(Dh(p)\big)$.

(3) Suppose (Ω, g) is a local parametrization of $U \cap M$ such that $g(u) = p$. Then $T_pM = \big(Dg(u)\big)(\mathbb{R}^m)$.

Obviously, T_pM is an m-dimensional suspace of $\mathbb{R}^n$. For an m-dimensional C^k submanifold M in $\mathbb{R}^n$, tangent vectors can also be characterized by the following way: $v \in T_pM$ if and only if there is a C^k curve c on M such that $c(0) = p$ and $c'(0) = v$.

The proof of Proposition 6.7 and Proposition 6.8 is left as an exercise.

Suppose E is a subset of $\mathbb{R}^n$ and p is an arbitrary point in $\mathbb{R}^n$. We define the cone

$$T_p^mE = \{v \in \mathbb{R}^n; \Theta^{*m}\big(E \cap X(p,v,\varepsilon), p\big) > 0 \text{ for any } \varepsilon > 0\},$$

which is called the *m-dimensional approximate tangent cone of E at p* and whose elements are called *m-dimensional approximate tangent vectors of E at p*. Obviously, $T_p^m E \subset T_p E$. When m is clear from context, we simply say the approximate tangent cone of E at p and approximate tangent vectors of E at p.

It is easy to verify

$$T_p^m E = \bigcap_{\tilde{E} \subset \mathbb{R}^m} \{T_p \tilde{E}; \Theta^m(E \setminus \tilde{E}, p) = 0\}.$$

Remark 6.9. If C is any compact subset of $\mathbb{R}^n \setminus T_p^m E$ and $T = \{p + rv; r > 0, v \in C\}$, then $\Theta^m(E \cap T, p) = 0$. In fact, $\{a + v; v \in C\}$ can be covered by a finite family of sets $X(p, v, \varepsilon_v)$ with $\Theta^m(E \cap X(p, v, \varepsilon_v), p) = 0$, with ε_v depending on v, and this family also covers T.

Remark 6.9 asserts that the density of E at p in any closed cone away from the approximate tangent cone $T_p^m E$ is zero.

Example 6.10. Let E be the union of x_1-x_2 plane and the x_3-axis in $\mathbb{R}^3$. Then $T_0 E$ is E itself and $T_0^2 E$ is the x_1-x_2 plane.

Now we start to discuss tangent cones of countably rectifiable sets. We first prove a decomposition of countably rectifiable sets similar to Lemma 5.6.

Lemma 6.11. *Suppose E is a countably m-rectifiable subset of $\mathbb{R}^n$ and $t \in (1, \infty)$. Then there exist compact subsets $A_1, A_2, A_3, \cdots$ of $\mathbb{R}^m$ and Lipschitz maps $\psi_1, \psi_2, \psi_3, \cdots : \mathbb{R}^m \to \mathbb{R}^n$ such that $\psi_1(A_1), \psi_2(A_2), \psi_3(A_3), \cdots$ are disjoint subsets of E and that*

$$E = E_0 \cup \bigcup_{i=1}^{\infty} \psi_i(A_i),$$

with $\mathcal{H}^m(E_0) = 0$. Moreover, for each $i = 1, 2, \cdots$

$$Lip(\psi_i|_{A_i}) \le t, \quad \psi_i|_{A_i} \text{ is one-to-one}, \quad Lip\big((\psi_i|_{A_i})^{-1}\big) \le t,$$

and

$$t^{-1}|v| \le |D\psi_i(p)v| \le t|v| \quad \text{for any } p \in A_i, v \in \mathbb{R}^m.$$

Proof. By Definition 6.1, we write

$$E = E_0 \cup \bigcup_{i=1}^{\infty} f_i(K_i),$$

where E_0 is a subset of $\mathbb{R}^n$ with $\mathcal{H}^m(E_0) = 0$ and, for each $i = 1, 2, \cdots$, K_i is a Borel subset of $\mathbb{R}^m$ and $f_i : K_i \to \mathbb{R}^n$ is Lipschitz. By Corollary 5.7, we assume that $f_i|_{K_i}$ is one-to-one and $J_m f_i(x) > 0$ for any $x \in K_i$. Furthermore, we assume by Theorem 4.3 that f_i is Liptschitz from $\mathbb{R}^m$ into $\mathbb{R}^n$. For each K_i and f_i, we apply Lemma 5.6 to get a countable collection $\{E_{ij}\}_{j=1}^{\infty}$ of Borel sets of $\mathbb{R}^m$ and a sequence of symmetric nonsingular linear maps $T_{ij} : \mathbb{R}^m \to \mathbb{R}^m$ such that

$$K_i = \bigcup_{j=1}^{\infty} E_{ij},$$

and for each $j = 1, 2, \cdots,$

$$\mathrm{Lip}(f_i|_{E_{ij}} \circ T_{ij}^{-1}) \le t, \quad \mathrm{Lip}(T_{ij} \circ (f_i|_{E_{ij}})^{-1}) \le t,$$

and

$$t^{-1}|T_{ij}v| \le |Df_i(p) \cdot v| \le t|T_{ij}v| \quad \text{for any } p \in E_{ij},\ v \in \mathbb{R}^m.$$

Now we set $\psi_{ij} = f_i \circ T_{ij}^{-1}$ and $A_{ij} = T_{ij}(E_{ij})$. The result follows by a simple approximation by compact subsets. □

Theorem 6.12. *Suppose E is a countably m-rectifiable subset of $\mathbb{R}^n$. Then, for $\mathcal{H}^m$-almost all x in E,*

$$\Theta^m(E, x) = 1,$$

and $T_x^m E$ is an m-dimensional subspace of $\mathbb{R}^n$.

Proof. We first consider a special case. Suppose $E = \psi(A)$, where A is a subset of $\mathbb{R}^m$ and $\psi : \mathbb{R}^m \to \mathbb{R}^n$ is a map such that $\psi|_A$ is one-to-one and that $t \in (1, \infty)$ is a Lipschitz constant for $\psi|_A$ and $(\psi|_A)^{-1}$. Let $p \in A$ be a point such that A has $\mathcal{L}^m$-density 1 at $p \in A$ and ψ has an injective differential at p. Now we claim

$$t^{-2m} \le \Theta_*^m\big(\psi(A), \psi(p)\big) \le \Theta^{*m}\big(\psi(A), \psi(p)\big) \le t^{2m},$$

and

$$T_{\psi(p)}^m \psi(A) = T_{\psi(p)}\psi(A) = \mathrm{im}\ D\psi(p).$$

To prove this, we assume $p = 0$, $\psi(p) = 0$ and write $D\psi(p) = L$. First we note that

$$A \cap B_{\frac{r}{t}} \subset \psi^{-1}\big(\psi(A) \cap B_r\big), \quad \psi(A) \cap B_r \subset \psi(A \cap B_{tr}),$$

and hence

$$t^{-2m}\frac{\mathcal{L}^m(A \cap B_{r/t})}{\omega(m)(r/t)^m} \le \frac{\mathcal{H}^m(\psi(A) \cap B_r)}{\omega(m)r^m} \le t^{2m}\frac{\mathcal{L}^m(A \cap B_{tr})}{\omega(m)(tr)^m},$$

for any $r \in (0, \infty)$. Therefore, the first conclusion follows because A has $\mathcal{L}^m$-density 1 at 0. Next, it is easy to see that

$$T_0\big(\psi(A)\big) \subset \mathrm{im} L.$$

To prove the oppositive inclusion, we suppose $v \in \mathbb{R}^m$ and $\varepsilon > 0$. By choosing $\eta > 0$ and $\xi > 0$ with

$$(|v| + \eta) \cdot \xi + \|L\| \cdot \eta \le \varepsilon,$$

we have for any r sufficiently small

$$|\psi(x) - Lx| \le \xi|x| \quad \text{for any } x \in B_{r/t}^m,$$

where $B_{r/t}^m$ is the ball with radius r/t and center at the origin in $\mathbb{R}^m$. It implies

$$\psi\big(X(0, v, \eta) \cap B_{r/t}\big) \subset X\big(0, Lv, \varepsilon\big),$$

because, if $x \in \mathbb{R}^m$, $s > 0$, $|x| \le r/t$ and $|sx - v| < \eta$, then

$$\begin{aligned}|s\psi(x) - Lv| &\le s|\psi(x) - Lx| + |L(sx - v)| \\ &\le |sx|\xi + \|L\| \cdot |sx - v| < (|v| + \eta)\xi + \|L\|\eta \le \varepsilon.\end{aligned}$$

Therefore,

$$\frac{\mathcal{H}^m\big(\psi(A)\cap X(0,Lv,\varepsilon)\cap B_r\big)}{\omega(m)r^m}\geq t^{-2m}\frac{\mathcal{H}^m\big(A\cap X(0,v,\eta)\cap B_{r/t}\big)}{\omega(m)(r/t)^m},$$

which approaches $t^{-2m}\mathcal{L}^m(X(0,v,\eta)\cap B_1)/\omega(m)$ as r approaches 0. We conclude that

$$\Theta_*^m\big(\psi(A)\cap X(0,Lv,\varepsilon),0\big)>0,$$

and hence $Lv\in T_0^m(\psi(A))$. This finishes the proof of the claim.

Now we consider a general case. Assuming $t\in(1,\infty)$, we choose A_i and ψ_i as in Lemma 6.11. By Lemma 3.5 and Lemma 3.4, we have

$$\Theta^{*m}(E,x)\leq 1,\quad \Theta^m\big(E\setminus\psi_i(A_i),x\big)=0,$$

for $\mathcal{H}^m$-almost all $x\in\psi_i(A_i)$. Furthermore, Lemma 3.3, Theorem 4.5 and Lemma 5.5(iii) imply that, for $\mathcal{H}^m$-almost all points $x\in\psi_i(A_i)$, the hypotheses of the special case hold with $p=(\psi|_{A_i})^{-1}x$. We conclude that

$$t^{-2m}\leq\Theta_*^m(E,x)\leq\Theta^{*m}(E,x)\leq 1,$$
$$T_x^mE=\mathrm{im}D\psi_i(p),$$

at $\mathcal{H}^m$-almost all $x\in\psi_i(A_i)$, with $p=(\psi|_{A_i})^{-1}x$. □

To conclude this section, we give another characterization of countably rectifiable sets and discuss its applications.

Lemma 6.13. *An $\mathcal{H}^m$-measurable set $E\subset\mathbb{R}^n$ is countably m-rectifiable if and only if*

$$E\subset\bigcup_{j=0}^{\infty}N_j,$$

where $\mathcal{H}^m(N_0)=0$ and N_j is an m-dimensional C^1 submanifold of $\mathbb{R}^n$ for each $j\geq 1$.

Proof. The *if* part is essentially trivial and is omitted. The *only if* part is an easy consequence of Theorem 4.6 as follows.

Without loss of generality, we assume E has the form $E\subset f(\mathbb{R}^m)$ for a Lipschitz function $f:\mathbb{R}^m\to\mathbb{R}^n$. By Theorem 4.6, we choose C^1 functions $g_1,g_2,\cdots$ on $\mathbb{R}^m$, such that

$$f(\mathbb{R}^m)\subset E_0\cup\big(\bigcup_{i=1}^{\infty}g_i(\mathbb{R}^m)\big),$$

where $\mathcal{H}^m(E_0)=0$. Then set

$$N_0=E_0\cup\big(\bigcup_{i=1}^{\infty}g_i(C_i)\big),$$

where $C_i=\{x\in\mathbb{R}^m;\ J_mg_i(x)=0\}$ and J_mg_i denotes the Jacobian of g_i. By Theorem 5.3, $\mathcal{H}^m(g_i(C_i))=0$, and hence $\mathcal{H}^m(N_0)=0$. Now for each $x\in\mathbb{R}^m\setminus C_i$, we let $U_i(x)$ be an open subset of $\mathbb{R}^m\setminus C_i$ containing x such that $g_i|_{U_i(x)}$ is one to one. By the Inverse Function Theorem, such a $U_i(x)$ exists and $N_i(x)\equiv g_i(U_i(x))$ is an m-dimensional C^1

submanifold of $\mathbb{R}^n$. We choose a countable collection $x_1, x_2, \cdots,$ of $\mathbb{R}^m \setminus C_i$ such that $g_i(\mathbb{R}^m \setminus C_i) \subset \cup_{j=1}^{\infty} N_i(x_j)$. Hence, we have

$$f(\mathbb{R}^m) \subset N_0 \cup \big(\bigcup_{i,j=1}^{\infty} N_i(x_j) \big).$$

This finishes the proof. $\square$

Another way to express Lemma 6.13 is to write

$$E = E_0 \cup \big(\bigcup_{j=1}^{\infty} E_j \big),$$

where $\mathcal{H}^m(E_0) = 0$ and, for each $j \geq 1$, there exist an open subset $A_j \subset \mathbb{R}^m$ and a C^1 map $g_j : A_j \to \mathbb{R}^n$ such that $E_j \subset g_j(A_j)$ and that g_j is one-to-one on A_j and $Dg_j(x)$ is injective for any $x \in A_j$.

By Lemma 6.13, an $\mathcal{H}^m$-measurable set $E \subset \mathbb{R}^n$ is purely m-unrectifiable if and only if $\mathcal{H}^m(E \cap N) = 0$ for any m-dimensional C^1-manifold N of $\mathbb{R}^n$.

As an application of Lemma 6.13, we give another characterization of approximate tangent spaces of countably rectifiable sets. We show that approximate tangent spaces can be obtained by the weak convergence of suitable measures.

For any $\lambda > 0$ and $x \in \mathbb{R}^n$, we define $\eta_{x,\lambda} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\eta_{x,\lambda}(y) = \frac{y - x}{\lambda} \quad \text{for any } y \in \mathbb{R}^n.$$

Theorem 6.14. *Suppose E is a countably m-rectifiable subset in $\mathbb{R}^n$ with $0 < \mathcal{H}^m(E) < \infty$. Then for $\mathcal{H}^m$-almost every $x \in E$*

$$\lim_{\lambda \to 0} \int_{\eta_{x,\lambda}(E)} f(y) d\mathcal{H}^m(y) = \int_{T_x^m E} f(y) d\mathcal{H}^m(y) \quad \textit{for any } f \in C_0(\mathbb{R}^n). \tag{6.1}$$

By a change of variables $z = \lambda y + x$, (6.1) is equivalent to

$$\lim_{\lambda \to 0} \frac{1}{\lambda^m} \int_E f\big(\frac{z - x}{\lambda}\big) d\mathcal{H}^m(z) = \int_{T_x^m E} f(y) d\mathcal{H}^m(y) \quad \text{for any } f \in C_0(\mathbb{R}^n).$$

Proof. We first consider a special case when $E = \psi(A)$ for an open subset $A \subset \mathbb{R}^m$ and a C^1 map $\psi : A \to \mathbb{R}^n$ such that ψ is one-to-one on A and $D\psi(x)$ is injective for any $x \in A$. Hence E is a C^1 m-dimensional submanifold in $\mathbb{R}^n$ and we prove that classical tangent planes are approximate tangent planes. Consider an arbitrary $x \in E$ with $x = \psi(p)$ for a $p \in A$ and $f \in C_0(\mathbb{R}^n)$. Then we have by Theorem 5.8

$$\begin{aligned}
\text{L.H.S.} &= \frac{1}{\lambda^m} \int_{\psi(A)} f\big(\frac{z - x}{\lambda}\big) d\mathcal{H}^m(z) \\
&= \frac{1}{\lambda^m} \int_A f\big(\frac{\psi(y) - \psi(p)}{\lambda}\big) J_m\psi(y) d\mathcal{L}^m(y) \\
&= \int_{\eta_{p,\lambda}(A)} f\big(\frac{\psi(p + \lambda v) - \psi(p)}{\lambda}\big) J_m\psi(p + \lambda v) d\mathcal{L}^m(v),
\end{aligned}$$

where $J_m\psi$ is the Jacobian of ψ. We note that the integration is over a bounded domain for λ small in the last integral. Hence we may replace $\eta_{p,\lambda}(A)$ by $\mathbb{R}^m$. Letting $\lambda \to 0$ and using Theorem 5.8 again, we obtain

$$\begin{aligned}\text{L.H.S.} \to &\int_{\mathbb{R}^m} f\big(D\psi(p)\cdot v\big)J_m\psi(p)d\mathcal{L}^m(v)\\ = &\int_{D\psi(p)(\mathbb{R}^m)} f(y)d\mathcal{H}^m(y).\end{aligned}$$

We note that $D\psi(\psi^{-1}(x))(\mathbb{R}^m)$ is the approximate tangent cone of E at x.

Now, we consider the general case. By Lemma 6.13, we assume

$$E = E_0 \cup \big(\bigcup_{j=1}^{\infty} E_j\big),$$

where $\mathcal{H}^m(E_0) = 0$ and, for each $j \geq 1$, there exist an open subset $A_j \subset \mathbb{R}^m$ and a C^1 map $g_j : A_j \to \mathbb{R}^n$ such that $E_j \subset g_j(A_j)$ and that g_j is one-to-one in A_j and $Dg_j(x)$ is injective for any $x \in A_j$. Then for any $f \in C_0(\mathbb{R}^n)$ and any $j \geq 1$, we have

$$\begin{aligned}\int_{\eta_{x,\lambda}(E)} fd\mathcal{H}^m = &\int_{\eta_{x,\lambda}(E\setminus E_j)} fd\mathcal{H}^m - \int_{\eta_{x,\lambda}(g_j(A_j)\setminus E_j)} fd\mathcal{H}^m\\ &+ \int_{\eta_{x,\lambda}(g_j(E_j))} fd\mathcal{H}^m.\end{aligned}$$

We assume supp$f \subset B_R(0)$ for some $R > 0$. Then, we have

$$\begin{aligned}\Big|\int_{\eta_{x,\lambda}(E\setminus E_j)} fd\mathcal{H}^m\Big| = &\frac{1}{\lambda^m}\Big|\int_{E\setminus E_j} f\big(\frac{\cdot - x}{\lambda}\big)d\mathcal{H}^m\Big|\\ \leq &\sup|f|\cdot\frac{1}{\lambda^m}\mathcal{H}^m\big((E\setminus E_j)\cap B_{\lambda R}(0)\big).\end{aligned}$$

By Theorem 3.4, we obtain for $\mathcal{H}^m$ almost all $x \in E_j$

$$\int_{\eta_{x,\lambda}(E\setminus E_j)} fd\mathcal{H}^m \to 0 \quad \text{as } \lambda \to 0.$$

Similarly, we have

$$\int_{\eta_{x,\lambda}(g_j(A_j)\setminus E_j)} fd\mathcal{H}^m \to 0 \quad \text{as } \lambda \to 0.$$

By what we just proved, we obtain

$$\int_{\eta_{x,\lambda}(g_j(A_j))} fd\mathcal{H}^m \to \int_{Dg_j(g_j^{-1}(x))(\mathbb{R}^m)} fd\mathcal{H}^m \quad \text{as } \lambda \to 0.$$

Hence $Dg_j(g_j^{-1}(x))(\mathbb{R}^m)$ is the approximate tangent space of E for $\mathcal{H}^m$ almost all $x \in E_j$. □

Remark 6.9 and the density one property in Theorem 6.12 can also be obtained by choosing appropriate f in (6.1).

Corollary 6.15. *Suppose E is a countably m-rectifiable subset in $\mathbb{R}^n$ with $0 < \mathcal{H}^m(E) < \infty$. Then for $\mathcal{H}^m$-almost every point $x \in E$,*

$$\Theta^m(E, x) = 1, \tag{6.2}$$

and for any $\alpha \in (0, 1)$

$$\lim_{r\to 0} \frac{1}{r^m}\mathcal{H}^m\big(E \cap B_r(x) \setminus X_\alpha(T_p^m E, x)\big) = 0, \tag{6.3}$$

where

$$X_\alpha(T_p^m E, x) = \{y; dist(y - x, T_p^m E) < \alpha|y - x|\}.$$

Proof. To prove (6.2), we approximate the characteristic function of $B_\lambda(x) \subset \mathbb{R}^n$ by functions in $C_0(\mathbb{R}^n)$. We first choose

$$g_\delta(y) = \begin{cases} 1, & |y| < 1 \\ 0, & |y| \geq 1 + \delta. \end{cases}$$

Then, we have

$$\int_{\eta_{x,\lambda}(E)} g_\delta(y) d\mathcal{H}^m(y) = \frac{1}{\lambda^m}\int_E g_\delta(\frac{y-x}{\lambda}) d\mathcal{H}^m(y)$$
$$\geq \frac{1}{\lambda^m}\int_{E\cap B_\lambda(x)} g_\delta(\frac{y-x}{\lambda}) d\mathcal{H}^m(y) \geq \frac{1}{\lambda^m}\mathcal{H}^m\big(E \cap B_\lambda(x)\big).$$

Therefore, we obtain

$$\limsup_{\lambda\to 0} \frac{1}{\lambda^m}\mathcal{H}^m\big(E \cap B_\lambda(x)\big) \leq (1+\delta)^m \omega(m).$$

Next, by choosing

$$f_\delta(y) = \begin{cases} 1, & |y| < 1 - \delta \\ 0, & |y| \geq 1, \end{cases}$$

we have

$$\liminf_{\lambda\to 0} \frac{1}{\lambda^m}\mathcal{H}^m\big(E \cap B_\lambda(x)\big) \geq (1-\delta)^m \omega(m).$$

These give (6.2).

To prove (6.3), we approximate the characteristic function of $B_r(x) \setminus X_\alpha(V_x, x)$ by functions in $C_0(\mathbb{R}^n)$. Essentially the same argument as in the proof of (6.2) yields (6.3). □

Remark 6.16. As we see in the next section, approximate tangent spaces can also be characterized by (6.3) with positive lower densities.

7. Weak Approximate Tangent Space Property

Throughout this section, m and n are two positive integers with $1 \leq m \leq n-1$. We first introduce some notations which will be used repeatedly.

Let $GL(n,m)$ be the Grassmannian manifold of all m-dimensional linear subspaces of $\mathbb{R}^n$. For any $V \in GL(n,m)$, $x \in \mathbb{R}^n$, $r \in (0,\infty)$ and $\alpha \in (0,1)$, a cone around the plane $V+x$ with the opening α is defined by

$$X_\alpha(V,x) = \{y; \mathrm{dist}(y-x, V) < \alpha|y-x|\},$$
$$X_\alpha(V,x,r) = \{y; \mathrm{dist}(y-x, V) < \alpha|y-x|, |y-x| \leq r\}.$$

Finally, for any subset $A \subset \mathbb{R}^n$, we write for $r > 0$,

$$A(r) = \{x \in \mathbb{R}^n;\ \mathrm{dist}(x, A) < r\}.$$

In this section, we discuss the rectifiability by approximate tangent spaces. We first prove a simple result.

Lemma 7.1. *Suppose $E \subset \mathbb{R}^n$ is an $\mathcal{H}^m$-measurable set such that for some $V \in GL(n,m)$, $\alpha \in (0,1)$ and $r \in (0,\infty)$,*

$$E \cap B_r(x) \subset X_\alpha(V,x) \quad \text{for any } x \in E.$$

Then E is countably m-rectifiable.

Proof. Without loss of generality, we assume $\mathrm{diam}(E) < r$, so that $E \subset B_r(x)$ for any $x \in E$. Let P_V be the orthogonal projection of $\mathbb{R}^n$ onto V. Then we have

$$|P_V x - P_V y| \geq \sqrt{1-\alpha^2}|x-y| \quad \text{for any } x, y \in E.$$

Hence $P_V|_E$ is injective. Taking $f = (P_V|_E)^{-1}$, we see that $f : P_V(E)(\subset V) \to \mathbb{R}^n$ is a Lipschitz function with $\mathrm{Lip}(f) \leq (1-\alpha^2)^{-1/2}$ and $E = f(P_V E)$. Thus, E is countably m-rectifiable. □

In Lemma 7.1, the m-dimensional subspace $V \in GL(n,m)$ is fixed for the entire set E. In fact, we may allow V, as well as α and r, to change according to x.

We now discuss the rectifiability by approximate tangent spaces.

Theorem 7.2. *Suppose E is an $\mathcal{H}^m$-measurable subset of $\mathbb{R}^n$ with $0 < \mathcal{H}^m(E) < \infty$. Suppose $\varepsilon_0 \in (0,1)$ and $\alpha \in (0,1/2)$ are two constants such that, for $\mathcal{H}^m$ almost every $x \in E$, there exists a $V = V_x \in GL(n,m)$ satisfying*

$$\Theta^m_*(E,x) \geq \varepsilon_0,$$

and

$$\lim_{r\to 0+} \frac{1}{r^m}\mathcal{H}^m\big(E \cap B_r(x) \setminus X_\alpha(V,x)\big) = 0.$$

Then E is countably m-rectifiable.

Proof. First, we set

$$f_i(x) = \inf_{0<r<\frac{1}{i}} \frac{1}{\omega(m)r^m}\mathcal{H}^m\big(E \cap B_r(x)\big),$$
$$g_i(x) = \sup_{0<r<\frac{1}{i}} \frac{1}{\omega(m)r^m}\mathcal{H}^m\big(E \cap B_r(x) \setminus X_\alpha(V_x, x)\big).$$

Then $f_i(x) \to \varepsilon_1$ and $g_i(x) \to 0$ for $\mathcal{H}^m$ almost all $x \in E$ as $i \to +\infty$, for a constant $\varepsilon_1 \geq \varepsilon_0$. By Egoroff's Theorem, for any $\epsilon > 0$, there is a closed subset $F \subset E$ such that $\mathcal{H}^m(E \setminus F) < \epsilon$ and that $f_i(x)$ and $g_i(x)$ converge uniformly on F. Hence, there exists a constant $\delta = \delta(\epsilon)$ such that

$$\frac{1}{\omega(m)r^m}\mathcal{H}^m\big(E \cap B_r(x)\big) \geq \varepsilon_0 - \epsilon, \tag{7.1}$$

and

$$\frac{1}{\omega(m)r^m}\mathcal{H}^m(E \cap B_r(x) \setminus X_\alpha(V_x, x)\big) \leq \epsilon, \tag{7.2}$$

for any $x \in F$ and $r \in (0, \delta]$.

Now, we choose a finite collection $\{V_j\}_{j=1}^N$ in $GL(n, m)$ such that, for any $V \in GL(n, m)$, there is a $1 \leq j \leq N$ with

$$\operatorname{dist}(V_j, V) = \|P_{V_j} - P_V\| < \frac{\alpha}{8}.$$

Let $F_j = \{x \in F; \operatorname{dist}(V_x, V_j) < \alpha/8\}$. Then

$$F = \bigcup_{j=1}^N F_j.$$

Now we claim

$$F_j \cap B_{\frac{\delta}{2}}(x) \subset X_{\frac{3\alpha}{2}}(V_j, x) \quad \text{for any } x \in F_j,\ j = 1, 2, \cdots, N. \tag{7.3}$$

If (7.3) were not true, we could find a point $x \in F_j$ and a point $y \in F_j \cap \partial B_r(x) \setminus X_{3\alpha/2}(V_j, x)$ for an $r \in (0, \delta/2]$. Since $x \in F$ and $2r \leq \delta$, we have by (7.2)

$$\mathcal{H}^m\big(E \cap B_{2r}(x) \setminus X_\alpha(V_x, x)\big) < \epsilon\omega(m)(2r)^m.$$

Meanwhile, a simple geometry shows that $B_{\alpha r/4}(y) \subset B_{2r}(x) \setminus X_\alpha(V_x, x)$.
In fact, we have

$$|z - x| \leq |y - x| + |z - y| \leq \big(1 + \frac{\alpha}{4}\big)r \leq \frac{9}{8}r,$$

and

$$\begin{aligned}|P_{V^\perp}(z - x)| \geq& |P_{V_j^\perp}(y - x)| - |(P_{V_j^\perp} - P_{V^\perp})(y - x)| - |P_{V^\perp}(z - x)|\\ \geq& \frac{3}{2}\alpha r - \frac{1}{8}\alpha r - \frac{1}{4}\alpha r = \frac{9}{8}\alpha r,\end{aligned}$$

where $P_{V^\perp}$ and $P_{V_j^\perp}$ are orthogonal projections onto $V^\perp$ and $V_j^\perp$ respectively. Then we obtain by (7.1)

$$\mathcal{H}^m\big(E\cap B_{2r}(x)\setminus X_\alpha(V_x,x)\big)\ge \mathcal{H}^m\big(E\cap B_{\frac{\alpha r}{4}}(y)\big)\ge \omega(m)\big(\frac{\alpha r}{4}\big)^m(\varepsilon_0-\epsilon).$$

It follows $\epsilon\ge(\alpha/8)^m(\varepsilon_0-\epsilon)$. This yields a contradiction if $\epsilon>0$ is chosen small, and hence proves (7.3). Now we apply Lemma 7.1 to conclude that F_j is countably m-rectifiable for each $j=1,\cdots,N$.

In conclusion, there exists a positive number $\epsilon(m,\alpha,\varepsilon_0)$ such that, for each $\epsilon<\epsilon(m,\alpha,\varepsilon_0)$, there exists a countably m-rectifiable subset $E_\epsilon\subset E$ such that

$$\mathcal{H}^m(E\setminus E_\epsilon)<\epsilon.$$

This implies that E itself is countably m-rectifiable. □

Theorem 7.3. *Suppose E is an $\mathcal{H}^m$-measurable subset of $\mathbb{R}^n$ with $0<\mathcal{H}^m(E)<\infty$. If for $\mathcal{H}^m$ almost every $x\in E$ there exists a m-dimensional hyperplane $V\in GL(n,m)$ satisfying*

$$\lim_{\lambda\to 0}\int_{\eta_{x,\lambda}(E)} f(y)d\mathcal{H}^m(y)=\int_V f(y)d\mathcal{H}^m(y)\quad \text{for any } f\in C_0(\mathbb{R}^n), \tag{7.4}$$

then E is countably m-rectifiable.

Proof. By Corollary 6.15, (7.4) implies (6.2) and (6.3). Hence Theorem 7.2 yields the desired result. □

A similar argument yields the following more general form.

Theorem 7.4. *Suppose μ is a Radon measure on $\mathbb{R}^n$, and for any $x\in\mathbb{R}^n$ and $\lambda>0$, let $\mu_{x,\lambda}$ be the measure given by $\mu_{x,\lambda}\ (A)=\lambda^{-m}\mu(x+\lambda A)$ for any subset $A\subset\mathbb{R}^n$. Suppose that, for μ almost all x, there is a $\theta(x)\in(0,\infty)$ and an m-dimensional subspace $V\in GL(n,m)$ with*

$$\lim_{\lambda\to 0}\int f(y)\theta(x+\lambda y)d\mu_{x,\lambda}(y)=\theta(x)\int_V f(y)d\mathcal{H}^m(y).$$

Let $E=\{x;$ the above identity holds for a $V\in GL(n,m)$ and a $\theta(x)\in(0,\infty)\}$, and set $\theta(x)=0$ for $x\in\mathbb{R}^n\setminus E$. Then E is countably m-rectifiable, θ is $\mathcal{H}^m$-measurable on E and $\mu=\mathcal{H}^m\lfloor\theta$.

In above results, we discuss the approximate tangent spaces by using cones. We may also use strips.

Corollary 7.5. *Suppose E is an $\mathcal{H}^m$-measurable subset of $\mathbb{R}^n$ with $0<\mathcal{H}^m(E)<\infty$. Suppose $\varepsilon_0\in(0,1)$ and $\eta\in(0,1/4)$ are two constants such that, for $\mathcal{H}^m$ almost every point $x\in E$, there exists an m-dimensional affine space $W=W_x$ of $\mathbb{R}^n$ containing x satisfying*

$$\Theta^m_*(E,x)\ge\varepsilon_0,$$

and

$$\lim_{r\to 0+} \frac{1}{r^m}\mathcal{H}^m\big(E\cap B_r(x)\setminus W(\eta r)\big) = 0. \tag{7.5}$$

Then E is countably m-rectifiable if and only if

Recall that $W(\beta) = \{y \in \mathbb{R}^n; \operatorname{dist}(y, W) < \beta\}$ for any positive number β.

Proof. Take an $x \in E$ such that (4.2.9) holds. Set $V = W - \{x\}$. We prove

$$\lim_{r\to 0+} \frac{1}{r^m}\mathcal{H}^m\big(E\cap B_r(x)\setminus X_{2\eta}(V,x)\big) = 0. \tag{7.6}$$

For any fixed $r > 0$, consider $y \in E \cap B_r(x) \setminus X_{2\eta}(V, x)$. Then there exists a positive integer i such that $r/2^i < |y - x| \le r/2^{i-1}$. It follows that $\operatorname{dist}(y - x, V) \ge 2\eta|y - x| > \eta r/2^{i-1}$, and hence $y \in E \cap B(x, r/2^{i-1}) \setminus W(\eta r/2^{i-1})$. Therefore we have

$$E\cap B_r(x)\setminus X_{2\eta}(V,x) \subset \bigcup_{i=1}^{\infty}\left(E\cap B_{r/2^{i-1}}(x)\setminus W(\eta r/2^{i-1})\right).$$

Now (7.6) follows from (7.5) easily. □

In the following, $A(x, n, m)$ denotes the collection of all m-dimensional affine spaces of $\mathbb{R}^n$ containing x.

To motivate our discussion, we first examine Corollary 7.5 and note that (7.5) can be formulated as follows. For any $\sigma > 0$, there exists an $r_0 > 0$ such that for any $r \in (0, r_0)$

$$\frac{1}{r^m}\mathcal{H}^m\big(E\cap B_r(x)\setminus W(\eta r)\big) < \sigma.$$

Here the affine subspace W depends only on the point x, independent of r. In many cases, W is allowed to change according to r. We still intend to conclude the rectifiability. To do this, we need to strengthen the hypothesis on the density lower bound.

Definition 7.6. A subset $E \subset \mathbb{R}^n$ has the weak approximate tangent space property at $x \in E$ if, for any positive constants σ and η, there exist positive contants r_0 and λ such that, for any $r \in (0, r_0)$, there exists an affine m-plane $W \in A(x, n, m)$ satisfying

$$\mathcal{H}^m\big(E\cap B_{\eta r}(y)\big) \ge \lambda r^m \quad \text{for any } y \in W\cap B_r(x), \tag{7.7}$$

and

$$\mathcal{H}^m\big(E\cap B_r(x)\setminus W(\eta r)\big) < \sigma r^m. \tag{7.8}$$

Such a W is called a weak approximate tangent space of E at x at the scale r.

In Definition 7.6, the affine m-plane W is allowed to depend on r.

Remark 7.7. According to (7.8), most of E lies near W in $B_r(x)$. Clearly, (7.7) implies the positiveness of the lower density of E, i.e., $\Theta^m_*(E, x) > 0$ for $\mathcal{H}^m$ almost all $x \in E$. Note that (7.7) also implies $W \cap B_r(x) \subset E(\eta r)$. Hence, there are no big holes in E near $W \cap B_r(x)$.

The main result is the following theorem.

Theorem 7.8. *Suppose E is an $\mathcal{H}^m$ measurable subset of $\mathbb{R}^n$ with $\mathcal{H}^m(E) < \infty$. If E satisfies the weak approximate tangent space property at $\mathcal{H}^m$ almost all points in E, then E is countably m-rectifiable.*

We first prove a key result on the projection.

Lemma 7.9. *Suppose E is an $\mathcal{H}^m$ measurable subset of $\mathbb{R}^n$ with $\mathcal{H}^m(E) < \infty$. If E satisfies the weak approximate tangent space property at $\mathcal{H}^m$ almost all points in E, then for $\mathcal{H}^m$ almost all $x \in E$*

$$\lim_{r\to 0} \sup_{V\in A(x,n,m)} \frac{1}{\omega(m)r^m}\mathcal{H}^m\big(P_V(E\cap B_r(x))\big) = 1. \tag{7.9}$$

We first discuss the idea of the proof. For simplicity, let E satisfy, for some positive constants c and d, $cr^m \le \mathcal{H}^m(E\cap B_r(x)) \le dr^m$ for any $x \in E$ and $r \in (0, r_0)$ and have the weak approximate tangent space property in a uniform manner. Consider a fixed ball $B_r(x)$ such that $E \cap B_r(x)$ lies close to an m-plane V. We need to prove that the projection of $E\cap B_r(x)$ on V cannot be too small. If otherwise, we find many disjoint open cylinders C_i of radii $\rho_i << r$ and orthogonal to V such that the cylinders with the same centers and radii $5\rho_i$ are disjoint, that $E \cap B_r(x) \cap C_i = \emptyset$ and that $B_r(x)\cap\partial C_i$ contains a point z_i of E. Then for a large constant M, $E\cap B_{M\rho_i}(z_i)$ is well approximated by an m-plane W_i. For most indices i, there is very little of E in $B_r(x)\cap C_i$, whence W_i must be almost orthogonal to V. This will yield many disjoint balls $B_{\rho_i}(x_{i,j}) \subset B_r(x)$ with $x_{i,j} \in E$ so that $\mathcal{H}^m(E \cap \cup_{i,j} B_{\rho_i}(x_{i,j}))$ will be much greater than r^m, which is a contradiction.

Proof of Lemma 7.9. Let $\varepsilon > 0$. Since E has positive lower densities at $\mathcal{H}^m$ almost all points E, there exist a compact subset F of E with $\mathcal{H}^m(E\backslash F) < \varepsilon$ and positive constants δ and r_0 such that for any $x \in F$ and $r \in (0, r_0)$,

$$\mathcal{H}^m\big(E \cap B_r(x)\big) > \delta\omega(m)r^m. \tag{7.10}$$

See Remark 7.7. We claim that, for any $\eta \in (0, \delta\varepsilon)$, we can find a compact subset $F_1 \subset F$ and a positive constant $r_1 \le r_0$ such that $\mathcal{H}^m(F \setminus F_1) < \varepsilon$ and that, for any $x \in F_1$ and $r \in (0, r_1)$, there is a $W \in A(x, n, m)$ satisfying

$$F \cap B_r(x) \setminus W(\eta r) = \emptyset, \tag{7.11}$$

and

$$W \cap B_r(x) \subset F(\eta r). \tag{7.12}$$

Note that (7.12) is equivalent to the relation

$$B_{\eta r}(y) \cap F \ne \emptyset \quad \text{for any } y \in W \cap B_r(x).$$

We first prove (7.11). By taking $\eta < \delta\varepsilon$ and $\sigma = \delta(\eta/4)^m$ in Definition 7.6, we find a compact subset $F_{11} \subset F$ and positive constant $r_{11} \le r_0$, with $\mathcal{H}^m(F \setminus F_{11}) < \varepsilon/2$, such that for any $x \in F_{11}$ and $r \in (0, r_{11})$ there is a $W \in A(x, n, m)$ satisfying

$$\mathcal{H}^m\big(E \cap B_r(x) \setminus W(\frac{\eta r}{4})\big) < \delta\big(\frac{\eta}{4}\big)^m r^m.$$

Then (7.11) holds for any $x \in F_{11}$ and $r \le r_{11}/2$. If not, we take a $y \in F \cap B_r(x) \setminus W(\eta r)$. On one hand, $y \in B_r(x) \setminus W(\eta r)$ implies $B_{\eta r/2}(y) \subset B_{2r}(x) \setminus W(\eta r/2)$. Hence, we have

$$\mathcal{H}^m\big(E \cap B_{\frac{\eta r}{2}}(y)\big) \le \mathcal{H}^m\big(E \cap B_{2r}(x) \setminus W(\frac{\eta r}{2})\big) < \delta\big(\frac{\eta r}{2}\big)^m.$$

On the other hand, by $y \in F$, (7.10) implies

$$\mathcal{H}^m\big(E \cap B_{\frac{\eta r}{2}}(y)\big) > \delta\big(\frac{\eta r}{2}\big)^m.$$

This is a contradiction. To prove (7.12), we first observe that it holds with F replaced by E by (7.7). Note that, when η is fixed, $\lambda = \lambda(x)$ in (7.7) depends only on x. Hence, we first take a compact subset $F' \subset F$, with $\mathcal{H}^m(F \setminus F') < \varepsilon/4$, such that $\lambda(x) \ge \lambda_0 > 0$ for any $x \in F'$, namely, for any $x \in F'$ and $r \in (0, r_0)$

$$\mathcal{H}^m\big(E \cap B_{\eta r}(y)\big) \ge \lambda_0 r^m \quad \text{for any } y \in W \cap B_r(x).$$

By Theorem 3.4, $\Theta^m(E \setminus F, x) = 0$ for $\mathcal{H}^m$ almost all $x \in F$. Hence, we have a compact set $F_{12} \subset F'$, with $\mathcal{H}^m(F' \setminus F_{12}) < \varepsilon/4$, and a constant $r_{12} \le r_0$ such that for any $x \in F_{12}$ and $r \le r_{12}$

$$\mathcal{H}^m\big((E \setminus F) \cap B_{2r}(x)\big) \le \frac{1}{2}\lambda_0 r^m.$$

This implies for any $x \in F_{12}$ and $r \le r_{12}$

$$\mathcal{H}^m\big(F \cap B_{\eta r}(y)\big) \ge \frac{1}{2}\lambda_0 r^m \quad \text{for any } y \in W \cap B_r(x).$$

and hence (7.12) holds for any $x \in F_{12}$ and $r \le r_{12}$. Now we set $F_1 = F_{11} \cap F_{12}$ and $r_1 = \min\{r_{11}/2, r_{12}\}$.

By Theorem 3.4 and Theorem 3.5, $\Theta^{*m}(E, x) \le 1$ and $\Theta^m(E \setminus F_1, x) = 0$ for $\mathcal{H}^m$ almost all $x \in F_1$. Hence, as before, we see that for $\mathcal{H}^m$ almost all $x \in F_1$ there exists a positive number $r_2 \le r_1$ such that for any $r \in (0, r_2)$ there is an affine m-space $W \in A(x, n, m)$ for which

$$F_1 \cap B_r(x) \setminus W(\eta r) = \emptyset, \tag{7.13}$$

$$W \cap B_r(x) \subset F_1(\eta r), \tag{7.14}$$

$$\mathcal{H}^m\big(E \cap B_r(x)\big) < 2\omega(m) r^m, \tag{7.15}$$

$$\mathcal{H}^m\big((E \setminus F_1) \cap B_r(x)\big) < 200^{-m}\omega(m) t \delta r^m, \tag{7.16}$$

for a small t to be fixed. We note that r_2 depends on $x \in F_1$. Compare with r_1, which depends only on F_1.

Fix such x, r and W. For simplicity, we assume $x = 0$ and $V = W \in GL(n, m)$. Let P_V be the orthogonal projection of $\mathbb{R}^n$ onto V. We prove that if, for given δ and $s \in (1/2, 1)$, the constant $\eta < (1 - s)/8$ is sufficiently small, then

$$\mathcal{H}^m\big(P_V(E \cap B_r(0)\big) \ge \omega(m) s^{2m} r^m. \tag{7.17}$$

If (7.17) holds, then

$$\limsup_{r \to 0} \sup_{V \in GL(n,m)} \frac{1}{\omega(m) r^m} \mathcal{H}^m\big(P_V(E \cap B_r(0)\big) \ge s^{2m}.$$

Since $s<1$ is arbitrary, we conclude (7.9) for $x=0$ as $\Theta^{*m}(E,x)\leq 1$ for $\mathcal{H}^m$ almost all $x\in E$. (Without evoking the rectifiability, we conclude from (7.17) that $\Theta^m(E,x)=1$, as $\Theta^{*m}(E,x)\leq 1$ for $\mathcal{H}^m$ almost all $x\in E$.)

Suppose (7.17) fails. Set

$$C=P_V\big(F_1\cap B_r(0)\big),\quad D=V\cap B_{sr}(0)\setminus C.$$

Then C is compact and

$$\mathcal{H}^m(C)<\omega(m)s^{2m}r^m.$$

Therefore,

$$\mathcal{H}^m(D)\geq\omega(m)(s^m-s^{2m})r^m=\omega(m)tr^m, \tag{7.18}$$

with $t=s^m-s^{2m}$. We emphasize that the precise value of $\mathcal{H}^m\big(V\cap B_{sr}(0)\big)$ is used, as $\mathcal{H}^m=\mathcal{L}^m$ on $\mathbb{R}^m$ by Theorem 2.10. We cover D by balls $B_\rho(x)$, with $x\in D$ and

$$C\cap B_\rho(x)=\emptyset,\quad C\cap\partial B_\rho(x)\neq\emptyset.$$

This is equivalent to

$$F_1\cap B_r(0)\cap P_V^{-1}\big(B_\rho(x)\big)=\emptyset,\quad F_1\cap B_r(0)\cap\partial P_V^{-1}\big(B_\rho(x)\big)\neq\emptyset.$$

By (7.14) and the inequality $\eta<1-s$, we obtain

$$\rho\leq\eta r.$$

Applying Theorem 1.5, the Vitali's covering theorem, to the balls $B_{5\rho}(x)$ as a covering of $\bar{D}$, we select a finite number of them, say $B_{5\rho_i}(x_i)$, $i=1,\cdots,l$, such that

$$B_{5\rho_i}(x_i)\cap B_{5\rho_j}(x_j)=\emptyset\quad\text{for } i\neq j, \tag{7.19}$$

and the balls $B_{25\rho_i}(x_i)$ cover D, hence by (7.18)

$$\sum_{i=1}^{l}\rho_i^m\geq 25^{-m}tr^m. \tag{7.20}$$

Note also

$$\rho_i\leq\eta r,\quad\text{for } i=1,\cdots,l. \tag{7.21}$$

For an $\alpha\in(0,1)$ to be determined, we consider the sets

$$S_i=P_V^{-1}\big(B_{\rho_i/2}(x_i)\big)\cap V(\alpha r),\quad\text{for any } i=1,\cdots,l,$$

and rearrange them so that they contain no points of F for $i=1,\cdots,k$ and that they contain at least one point, say y_i of F, for $i=k+1,\cdots,l$.
By $x_i\in V\cap B_{sr}(0)$ and (7.21) for $i=k+1,\cdots,l$, we have

$$\begin{aligned}|y_i|&\leq|x_i|+|x_i-P_Vy_i|+|P_Vy_i-y_i|\\&\leq sr+\frac{1}{2}\rho_i+\alpha r\leq sr+\frac{1}{2}\eta r+\alpha r.\end{aligned}$$

Hence $B_{\rho_i/4}(y_i)\subset B_r(0)$ provided

$$sr+\frac{1}{2}\eta r+\alpha r+\frac{1}{4}\eta r\leq r.$$

Therefore, we take $\alpha = (1-s)/4$ and $\eta < 1-s$. Moreover,

$$P_V\big(B_{\frac{\rho_i}{4}}(y_i)\big) \subset V \cap B_{\rho_i}(x_i) \subset V \setminus C. \tag{7.22}$$

It follows that

$$\bigcup_{i=k+1}^{l} E \cap B_{\frac{\rho_i}{4}}(y_i) \subset (E \setminus F_1) \cap B_r(0). \tag{7.23}$$

Since the balls $B_{\rho_i/4}(y_i)$ are disjoint by (7.22) and (7.19), we obtain from (7.10), (7.23) and (7.16),

$$\omega(m)\delta 4^{-m} \sum_{i=k+1}^{l} \rho_i^m < \omega(m) 200^{-m} t\delta r^m,$$

and hence by (7.20),

$$\sum_{i=1}^{k} \rho_i^m > 50^{-m} t r^m. \tag{7.24}$$

From now on, we only consider $i = 1, \cdots, k$. Recalling (7.13) and how the balls $B_{\rho_i}(x_i)$ were chosen, we see that there are points

$$z_i \in P_V^{-1}\big(\partial B_{\rho_i}(x_i)\big) \cap V(\eta r) \cap F_1 \quad \text{for } i = 1, \cdots, k. \tag{7.25}$$

By (7.21), $\eta^{-1}\rho_i \le r < r_1$, and we can apply (7.12) to z_i and $\eta^{-1}(1-s)\rho_i/16$ to get $W_i \in A(z_i, n, m)$ such that

$$A_i \equiv B_{\eta^{-1}(1-s)\rho_i/16}(z_i) \cap W_i \subset F\big(\frac{1}{16}(1-s)\rho_i\big). \tag{7.26}$$

Then $x_i \notin P_V A_i$. In fact, if there were $x \in A_i$ with $P_V x = x_i$, we could find by (7.26) a point $y \in F$ with $|x-y| \le (1-s)\rho_i/16$. Then $P_V y \in B_{\rho_i/2}(x_i)$ and by (7.25), (7.21) and $\eta < (1-s)/8$,

$$\begin{aligned} d(y,V) \le& |y-x| + |x-z_i| + d(z_i, V) \\ \le& \frac{1}{16}(1-s)\rho_i + \frac{1}{16}\eta^{-1}(1-s)\rho_i + \eta r < \frac{1}{4}(1-s)r, \end{aligned}$$

and hence y would belong to $S_i \cap F$, which is empty. We note that the fact $x_i \notin P_V A_i$ illustrates that W_i is almost orthogonal to V.

Let I_i be the closed line segment with endpoints x_i and $P_V z_i$. Then $I_i \cap \partial_V P_V(A_i) \neq \emptyset$, since $x_i \in I_i \setminus P_V(A_i)$ and $P_V z_i \in I_i \cap P_V A_i$. Here ∂_V means the boundary relative to V. Since $\partial_V P_V(A_i) = P_V(\partial_{W_i} A_i)$, we find $w_i \in \partial_{W_i} A_i$ such that $P_V w_i \in I_i$. Let J_i be the closed line segment connecting z_i and w_i. Then $J_i \subset A_i$ and $P_V J_i \subset I_i$, hence by (7.25)

$$|P_V x - x_i| \le \rho_i \quad \text{for any } x \in J_i. \tag{7.27}$$

By (7.26), J_i is contained in the union of the balls $B_{\rho_i}(x)$, $x \in F$. Since the length of J_i is $\eta^{-1}(1-s)\rho_i/16$, we choose a finite number of them, say $B_{\rho_i}(x_{i,j})$, $j = 1, \cdots, p$, such that,

$$J_i \cap B_{\rho_i}(x_{i,j}) \neq \emptyset, \tag{7.28}$$

$$B_{\rho_i}(x_{i,j}) \cap B_{\rho_i}(x_{i,q}) = \emptyset, \quad \text{for } j \neq q, \tag{7.29}$$

and

$$p > 160^{-1}\eta^{-1}(1-s). \tag{7.30}$$

Set for $i = 1, \cdots, k$

$$B_i = \bigcup_{j=1}^{p} B_{\rho_i}(x_{i,j}).$$

It follows from (7.27) and (7.28) that $P_V B_i \subset B_{3\rho_i}(x_i)$, whence by (4.3.18) the sets B_i are disjoint. By (7.28), (7.21) and $\eta < (1-s)/8$,

$$|x_{i,j} - w_i| \leq \mathcal{H}^1(J_i) + \rho_i = \frac{1}{16}\eta^{-1}(1-s)\rho_i + \rho_i < \frac{1}{4}(1-s)r.$$

With $x_i \in B_{sr}(0) \cap V$, as before, (7.25) and (7.21) yield

$$\begin{aligned}|w_i - x_i| \leq& |w_i - z_i| + |z_i - P_V z_i| + |P_W z_i - x_i| \\ \leq& \frac{1}{16}\eta^{-1}(1-s)\rho_i + \eta r + \rho_i \leq \frac{1}{16}(1-s)r + 2\eta r < \frac{1}{2}(1-s)r.\end{aligned}$$

It follows that $B_i \subset B_r(0)$. Using (7.29), (7.10) and (7.30), we obtain for $i = 1, \cdots, k$,

$$\begin{aligned}\mathcal{H}^m(E \cap B_i) =& \sum_{j=1}^{p} \mathcal{H}^m\big(E \cap B_{\rho_i}(x_{i,j})\big) \\ >& p\delta\omega(m)\rho_i^m > 160^{-1}(1-s)\eta^{-1}\delta\omega(m)\rho_i^m,\end{aligned}$$

and hence by (7.15) and (7.24),

$$\begin{aligned}2\omega(m)r^m > \mathcal{H}^m\big(E \cap B_r(0)\big) &\geq \sum_{i=1}^{k} \mathcal{H}^m(E \cap B_i) \\ &\geq 160^{-1}(1-s)\eta^{-1}\delta\omega(m)\sum_{i=1}^{k}\rho_i^m \\ &> 160^{-1}50^{-m}(1-s)\eta^{-1}\delta t\omega(m)r^m.\end{aligned}$$

As we may choose an arbitrary small η for given δ and s, we obtain a contradiction. □

Now, we begin to prove Theorem 7.8. The idea of the proof is quite simple. We prove by contradiction. Assuming E is purely m-unrectifiable, we prove $\mathcal{H}^m(P_V E) = 0$ for any $V \in GL(n, m)$. Take any $V \in GL(n, m)$. The weak tangent space property and the unrectifiablity imply that many tangent planes at small scales are almost verdict to V. Hence the projection of E onto V occupies a small portion. To be slightly more specific, consider any ball $B_r(x)$ with $x \in E$. The weak approximate tangent space property implies the existence of an affine m-plane W containing x such that the most part of E in $B_r(x)$ lies close to W. By the unrectifiability, a certain amount of E is around

$V^\perp$. Hence W is close to $V^\perp$. Therefore, most part of E in $B_r(x)$ projects into a small neighborhood of $P_V x$ in V. The last assertion is reflected in (7.34) in the proof below.

Proof of Theorem 7.8. We prove by a contradiction argument. By assuming E is purely m-unrectifiable, we prove $\mathcal{H}^m(P_V E) = 0$ for any $V \in GL(n,m)$. This clearly contradicts (7.9).

Let $\varepsilon \in (0,1/2)$ and $V \in GL(n,m)$. Then we find a compact subset $F \subset E$ and positive numbers r_0, δ and η, with $\mathcal{H}^m(E \setminus F) < \varepsilon$ and $\eta < \delta\varepsilon$, such that for any $x \in F$ and $r \in (0, r_0)$ there is a $W \in A(x,n,m)$ satisfying

$$\mathcal{H}^m\big(E \cap B_r(x)\big) > \delta r^m, \tag{7.31}$$

and

$$F \cap B_r(x) \setminus W(\eta r) = \emptyset. \tag{7.32}$$

The proof is similar to that of (7.10) and (7.11) in the proof of Lemma 7.9. Note that we also have

$$\mathcal{H}^m\big(P_V(E \setminus F)\big) < \varepsilon. \tag{7.33}$$

Now we claim that, for $\mathcal{H}^m$ almost all $x \in F$, there exists sufficiently small r such that

$$\mathcal{H}^m\big(P_V(F \cap B_r(x))\big) \le 10^m \eta r^m. \tag{7.34}$$

To prove this, we first note by Lemma 7.1 that $\{x \in F; F \cap X_\eta(V^\perp, x, 1/i) = \emptyset\}$ is countably m-rectifiable for any integer i. Since F is purely m-unrectifiable, then

$$\mathcal{H}^m\big(\bigcup_{i=1}^\infty \{x \in F; F \cap X_\eta(V^\perp, x, \frac{1}{i}) = \emptyset\}\big) = 0.$$

Hence for $\mathcal{H}^m$ almost all $x \in F$, there are points $y \in F$ arbitrarily close to x such that

$$|P_V(y-x)| < \eta|y-x|. \tag{7.35}$$

Let $x \in F$ be such a point and take a $y \in F$ such that $x, y \in F$ satisfy (7.35) and $r = |x-y| < r_0$. Let W be the weak tangent plane at x at the scale r as in (7.32). Intuitively, W is close to $V^\perp$. Hence the projection of W, or even $W(\eta r)$, which contains F, into V should be small. First by (7.32), we see $y \in W(\eta r)$. By setting $z = P_W y$, we have

$$|z-y| \le \eta r, \quad \frac{r}{2} \le |z-x| \le r, \quad |P_V(z-x)| < 2\eta r.$$

We select an orthonormal basis $\{e_1, \cdots, e_m\}$ for $W - \{x\}$ such that $P_V(e_i) \cdot P_V(e_j) = 0$ for $i \ne j$. Then for an $i \in \{1, \cdots, m\}$,

$$|P_V e_i| \le 2r^{-1}|P_V(z-x)| < 4\eta,$$

because otherwise we should have

$$\begin{aligned}|P_V(z-x)|^2 &= \sum_{j=1}^m |(z-x)\cdot e_j|^2 |P_V e_j|^2 \\ &> 4r^{-2}|P_V(z-x)|^2|z-x|^2 \ge |P_V(z-x)|^2.\end{aligned}$$

It follows that $P_V(W \cap B_r(x))$ is contained in an m-dimensional rectangle with one side of the length $8\eta r$ and the others of length $2r$. Hence by (7.32), $P_V(F \cap B_r(x))$ is contained in a rectangle with side lengths $10\eta r, 2r + 2\eta r, \cdots, 2r + 2\eta r$. Therefore, as $\eta < 1/2$ and $\mathcal{H}^m = \mathcal{L}^m$ on $\mathbb{R}^m$, we have (7.34).

Note that the collection of those balls $B_r(x)$, $x \in F$, is a fine cover of F. We use Corollary 1.8 to obtain disjoint $B_{r_i}(x_i)$ satisfying (7.34), $x_i \in F$, and

$$\mathcal{H}^m\big(F \setminus \bigcup_{i=1}^{\infty} B_{r_i}(x_i)\big) = 0.$$

Hence, we have

$$\mathcal{H}^m\big(P_V(F)\big) \leq \sum_{i=1}^{\infty} \mathcal{H}^m\big(P_V(F \cap B_{r_i}(x_i))\big) \leq 10^m \eta \sum_{i=1}^{\infty} r_i^m,$$

and by (7.31)

$$\mathcal{H}^m\big(P_V(F)\big) \leq 10^m \eta \delta^{-1} \sum_{i=1}^{\infty} \mathcal{H}^m\big(E \cap B_{r_i}(x_i)\big) \leq 10^m \varepsilon \mathcal{H}^m(E).$$

Combining with (7.33), we obtain $\mathcal{H}^m(P_V(E)) < (1 + 10^m \mathcal{H}^m(E))\varepsilon$. This holds for any ε small. □

References

[1] L. C. Evans, R. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press, Boca Raton, 1992.
[2] H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin-Heidelberg-New York, 1969.
[3] F. Lin, X. Yang, *Geometric Measure Theory - An Introduction*, International Press, Boston, 2002.
[4] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiablity*, Cambridge University Press, Cambridge, 1995.
[5] L. Simon, *Lectures on Geometric Measure Theory*, Proceedings of the Centre for Mathematical Analysi, Australian National University, Canberra, 1983.