Indian Statistical Institute, Kolkata

Mid-Semester Report in Geometric Measure Theory

An Exposition of the Area Formula

Author: Jishnu Parayil Shibu

Supervisor: Partha Sarathi Chakraborty

Date: October 5, 2025

Contents

0 Introduction

The aim of this report is to present and discuss the *generalised change of variables formula* in Geometric Measure Theory (GMT). This theorem, which follows from another central result known as the *area formula*, was first proved by Herbert Federer [4] and forms one of the cornerstones of GMT.

Before stating these two results, we begin by exploring the fundamental question the area formula seeks to answer.

Question 0.1. Given $A \subset \mathbb{R}^n$, can we relate the "n-dimensional volume" of the image of A under a Lipschitz map $f : \mathbb{R}^n \to \mathbb{R}^m$ where $n \leq m$ to something simpler - namely, the "first order approximation" of f at each point?

The area formula gives an affirmative answer. While it remains to make precise what we mean by the "n-dimensional volume" in $\mathbb{R}^m$ and by a "first–order approximation" (since f is not assumed to be differentiable everywhere), this question should not feel unfamiliar. Indeed, it extends a result already known for a narrower class of maps, namely C^1 functions. This is precisely the subject of the classical change of variables formula:

Theorem 0.1 (Classical Change of Variables)**.** *Let $U, V \subset \mathbb{R}^n$ be open, and let $f : U \to V$ be a C^1 diffeomorphism. Then, for every continuous function $h : V \to \mathbb{R}$,*

$$\int_V h(y)\, dy = \int_U h(f(x))\, |\det Df(x)|\, dx.$$

In particular, letting $h \equiv 1$ on V, we get $\mathrm{vol}(f(U)) = \int_U |\det Df(x)|\, dx$.

Thus, for C^1 maps and open sets A, the volume $f(A)$ depends only on the Jacobian Jf, which encodes the first order approximation of f. Motivated by this insight, we now state the area formula, and subsequently the generalised change of variables.

Theorem 0.2 (Area Formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz map with $n \leq m$, and let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$–measurable. Then*

$$\int_A Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \mathcal{H}^0\Big(A \cap f^{-1}(y)\Big)\, d\mathcal{H}^n(y),$$

Corollary 0.3 (Generalised Change of Variables)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz with $n \leq m$, and let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$–measurable. Then, for any $\mathcal{L}^n$–integrable function $g : A \to \mathbb{R}$,*

$$\int_A g(x) Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \sum_{x \in A \cap f^{-1}(y)} g(x)\, d\mathcal{H}^n(y),$$

Although a rigorous understanding of these results require substantial machinery, which we shall develop in subsequent sections, let's provide an intuitive picture and see how it relates to the classical change of variables as well as our fundamental question.

Note that $\mathcal{L}^n$ is the n-dimensional Lebesgue outer measure in $\mathbb{R}^n$, and the k–dimensional Hausdorff measure $\mathcal{H}^k$ in $\mathbb{R}^n$ (defined precisely in 2.1) is, informally, a measure of "k–dimensional volume". Later, in 3.24, we will show that the volume form of a C^1 k–dimensional submanifold M of $\mathbb{R}^n$ does coincide with the restriction of $\mathcal{H}^k$ on M . In particular, we shall see that $\mathcal{H}^n = \mathcal{L}^n$ in 2.24. Moreover, $\mathcal{H}^0$ is just the *counting measure*, so that $\mathcal{H}^0\Big(A \cap f^{-1}(y)\Big) = \#\Big(A \cap f^{-1}(y)\Big)$,

With these observations, notice that the RHS of 0.2 is nothing but the "n-dimensional volume" of $f(A)$, accounting for multiplicity. Taking f to be injective and $m = n$, the RHS is $\mathcal{L}^n(A)$ (as $H^0\Big(A \cap f^{-1}(y)\Big) = 1$ and $H^n = \mathcal{L}^n$).

Thus, in particular, if f is a C^1 diffeomorphism, we recover the volume formula obtained via the classical change of variables, by remarking that the Jacobian $Jf(x)$ in 0.2 is just the usual Jacobian in this case (we shall see that the more general definition of $Jf(x)$ for Lipschitz f does agree with this). The same is true for the classical change of variables formula which is a particular case of 0.3. To see this, take $g = h \circ f$ in 0.3.

Thus, the area formula is truly a powerful generalisation. To gain some insight into what $Jf(x)$ is, consider first a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$. Then $J(T)$ is a generalization of $|det(T)|$ (which is defined only when $m = n$) in the following sense:

$$J(T) = \frac{\mathcal{H}^n\Big(T(B)\Big)}{\mathcal{L}^n(B)}$$

where B is an $\mathcal{L}^n$-measurable set, i.e. $J(T)$ measures the "n–dimensional volume" distortion under T. That $J(T)$ does not depend on the choice of B is a nontrivial fact, which will be discussed in 3.1. When $m = n$, the classical change of variables formula indicates that this should coincide with $det(T)$.

Now, note that by Rademacher's Theorem (see 3.13), a Lipschitz function is differentiable $\mathcal{L}^n$ a.e. Thus $Jf(x)$ is defined to be just $J(Df(x))$. This should make the LHS of the Area Formula feel more intuitive.

With this intuition in mind, we now develop the theory leading to the area formula and the resulting change-of-variables theorem. Note that the primary reference for this report is [3], and other references, when used, shall be mentioned.

1 Preliminary Measure Theory

1.1 Radon Outer Measures

We begin by defining Radon outer measures and exploring some of their elementary properties. Unlike most treatments, we shall work almost exclusively with outer measures, as this allows us to handle even "non-measurable" sets. Although our later discussions will mostly involve $\mathbb{R}^n$ or, at most, metric spaces, we introduce the measure-theoretic prerequisites in their maximal generality, which will be useful for later applications. The references for this section are [2, 3].

Notation 1.1 (Outer measures and measurable sets)**.** Let X be a Hausdorff topological space and $\mu^*\colon \mathcal{P}(X) \to [0,\infty]$ an outer measure on X.

- Let $\mathcal{M}(X) := \{A \subseteq X : A \text{ is } \mu^*\text{-measurable}\}$
$$= \Big\{A \subseteq X : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subseteq X\Big\}$$
denote the σ-algebra of μ^*-measurable sets.

- The restriction of μ^* to $\mathcal{M}(X)$ is written $\mu := \mu^*|_{\mathcal{M}(X)}$.

  Thus, for $A \in \mathcal{M}(X)$ we write $\mu(A)$, while for sets $A \subseteq X$ that are not measurable, we continue to use $\mu^*(A)$. This convention will be useful, as only measurable sets enjoy countable additivity; for general sets, one only has countable subadditivity.

- The Borel σ-algebra of X is denoted $\mathcal{B}(X)$.

- For $C \subseteq X$, define the restriction outer measure of μ^* to C by $\mu^*\big|_C(A) := \mu^*(A \cap C), \quad A \subseteq X$.

Definition 1.1 (Regularity For Outer Measure)**.**

- We say that μ^* is *Borel* if $\mathcal{B}(X) \subseteq \mathcal{M}(X)$.

- Let $\Sigma \subseteq \mathcal{M}(X)$ be a σ-algebra. We say that μ^* is Σ-*regular* if for every $A \subseteq X$ there exists $S \in \Sigma$ with $A \subseteq S$ and $\mu^*(A) = \mu(S)$.

- In particular, μ^* is *Borel regular* if it is Borel and $\mathcal{B}(X)$-regular.

Proposition 1.2. *Let μ^* be a Borel regular outer measure on X and let $C \subseteq X$ be μ-measurable such that $\mu(C) < \infty$. Then $\mu^*\big|_C$ is also Borel regular.*

Proof. Since μ^* is Borel regular, every Borel set is μ^*-measurable. Hence for each Borel $B \subseteq X$ the intersection $B \cap C$ is μ^*-measurable, so $\mu^*\big|_C$ is a Borel outer measure.

Fix $A \subseteq X$. We need to find a set Borel set B s.t. $A \subseteq B$ and $\mu^*|_C(A) = \mu^*|_C(B)$, i.e. $\mu^*(A \cap C) = \mu(B \cap C)$. By regularity of μ^*, and since $\mu(C) < \infty$,

$$\exists B_1 \text{ Borel s.t. } C \subseteq B_1 \text{ and } \mu(B_1 \setminus C) = 0.$$

Apply Borel-regularity to the set $A \cap B_1$,

$$\exists B_2 \text{ Borel s.t. } A \cap B_1 \subseteq B_2 \text{ and } \mu^*(A \cap B_1) = \mu(B_2)$$

Let $B = B_2 \cup (X \setminus B_1)$. Then B is Borel and $A \subseteq B$: indeed, if $x \in A$ and $x \in B_1$ then $x \in A \cap B_1 \subseteq B_2$, while if $x \in A$ and $x \notin B_1$ then $x \in X \setminus B_1$.

Note that as $C \subseteq B_1$,

$$B \cap C = (B_2 \cap C) \cup ((X \setminus B_1) \cap C) = B_2 \cap C$$

Now $A \subseteq B \implies \mu^*(A \cap C) \leq \mu^*(B \cap C)$. Conversely, as $\mu(B_1 \setminus C) = 0$ ($\implies \mu^*(A \cap (B_1 \setminus C) = 0)$, we have:

$$\begin{aligned}
\mu^*(A \cap C) &= \mu^*(A \cap C) + \mu^*(A \cap (B_1 \setminus C)) \\
&\geq \mu^*(A \cap B_1) \\
&= \mu(B_2) \\
&\geq \mu(B_2 \cap C) = \mu(B \cap C)
\end{aligned}$$

Thus $\mu^*|_C(A) = \mu^*|_C(B)$, concluding the proof. □

Definition 1.3 (Inner and Outer Regularity)**.** Let X be a Hausdorff topological space, μ^* an outer measure on X, and let $S \subseteq \mathcal{P}(X)$.

- We say that μ^* is *outer regular on S* if for every $A \in S$,

$$\mu^*(A) = \inf\{\mu^*(U) : A \subseteq U,\ U \text{ open in } X\}.$$

- We say that μ^* is *inner regular on S* if for every $A \in S$,

$$\mu^*(A) = \sup\{\mu^*(K) : K \subseteq A,\ K \text{ compact in } X\}.$$

The same definitions apply when μ^* is a measure (not just an outer measure).

Definition 1.4 (Radon Outer Measure)**.** Let X be a Hausdorff topological space and μ^* an outer measure on X. We say that μ^* is a *Radon outer measure* if:

1. μ^* is Borel regular and locally finite, i.e. $\mu^*(K) < \infty$ for all compact $K \subseteq X$.

2. μ^* is outer regular on all subsets of X (even non-measurable subsets).

3. μ^* is inner regular on all open subsets of X.

Remark 1.5. In practice, when considering more specific kinds of spaces X, some of the conditions listed above could follow from one another. In particular, as we shall show in the remaining part of this section in Proposition 1.8, when X is a locally compact Hausdorff space that is also σ-compact (i.e. X can be written as a countable union of compact subsets), conditions (2) and (3) above follow automatically from (1).

We first prove a simple lemma which shall be useful in deriving these equivalences.

Lemma 1.6. *Let X be a locally compact, Hausdorff, σ-compact space. Then X admits a compact exhaustion, i.e. there exists an increasing sequence of compact sets $\{K_m\}_{m\geq 1}$ such that*

$$K_1 \subseteq \operatorname{int} K_2 \subseteq \operatorname{int} K_3 \subseteq \cdots, \qquad \bigcup_{m=1}^{\infty} K_m = X.$$

Proof. Since X is σ-compact, write $X = \bigcup_{m=1}^{\infty} D_m$ with each D_m compact. By local compactness, every $x \in D_m$ has an open neighbourhood U_x with compact closure $\overline{U_x}$. Because D_m is compact, finitely many of these neighbourhoods suffice; let L_m be the union of the corresponding compact closures. Then L_m is compact, and each point of D_m lies in the interior of some $\overline{U_x}$, hence $D_m \subseteq \operatorname{int} L_m$.
Now define

$$K_m := \bigcup_{j=1}^{m} L_j.$$

Each K_m is compact as a finite union of compact sets, and we have

$$K_m \subseteq \operatorname{int} K_{m+1},$$

since L_{m+1} already contains D_{m+1} in its interior. Finally, $\bigcup_{m=1}^{\infty} K_m = X$, as required. $\square$

We are now ready to prove regularity properties of simply Borel outer measures, i.e. outer measures for which Borel sets are measurable.

Proposition 1.7. *Let X be a locally compact Hausdorff space in which every open subset is σ-compact. Let μ be a Borel measure on X. Then*

(i) μ is inner regular on Borel sets with finite measure

(ii) If additionally μ is locally finite, then μ is outer regular on Borel sets.

Proof. **(i) Inner regularity on finite-measure Borel sets.** Let B be Borel s.t. $\mu(B) < \infty$. Set $\nu := \mu|_B$. Then ν is a finite Borel measure (since $\nu(X) = \mu(B) < \infty$). Define

$$\mathcal{F} := \{A \subseteq X :\ A \text{ is } \nu\text{-measurable and for each } \delta > 0,\ \exists \text{ compact } K \subseteq A \text{ with } \nu(A \backslash K) < \delta\}$$

Clearly, compact sets belong to $\mathcal{F}$ by definition. We begin by showing $\mathcal{F}$ is closed under countable union. Let $A_n \in \mathcal{F}$, $A = \bigcup_{n\geq 1} A_n$, and fix $\delta > 0$. For each n choose compact

$K_n \subseteq A_n$ with $\nu(A_n \setminus K_n) < \delta/2^{n+1}$. Define $K^{(m)} := \bigcup_{n=1}^m K_n$ (compact) and note that as $\nu(A) < \infty$, the decreasing sequence

$$\nu(A \setminus K^{(m)}) \downarrow \nu(\bigcup_{n\geq 1} A_n \setminus \bigcup_{n\geq 1} K_n) \leq \nu\Big(\bigcup_{n\geq 1}(A_n \setminus K_n)\Big) \leq \sum_{n\geq 1} \nu(A_n \setminus K_n) < \delta/2$$

Choosing m large, we can get $\nu(E_m) < \delta$, proving $A \in \mathcal{F}$. Closure under countable intersections can be proved analogously.

By hypothesis every open set U is σ-compact, and as compact sets belong to $\mathcal{F}$, $U \in \mathcal{F}$. Note also that by Lemma 1.6, we can choose a compact exhaustion $K_1 \subseteq K_2 \subseteq \cdots$ with $\bigcup_m K_m = X$ and $K_m \subseteq \operatorname{int} K_{m+1}$. Thus for any closed set E, the sets $E \cap K_n$ are compact, hence belong to $\mathcal{F}$, and $E = \bigcup_n (E \cap K_n)$, so $E \in \mathcal{F}$.

Let $\mathcal{G} := \{A \in \mathcal{F} : \ X \setminus A \in \mathcal{F}\}$. As both open and closed sets belong to $\mathcal{F}$, open sets in particular belong to $\mathcal{G}$. Note that $\mathcal{G}$ itself is closed under complements and countable unions: if $A \in \mathcal{G}$ then both A and $X \setminus A$ belong to $\mathcal{G}$ by definition. If $A_i \in \mathcal{G}$ for $i \geq 1$ then $\bigcup_i A_i \in \mathcal{F}$ (closure of $\mathcal{F}$ under countable unions) and $X \setminus \bigcup_i A_i = \bigcap_i (X \setminus A_i) \in \mathcal{F}$ (closure of $\mathcal{F}$ under countable intersections), so $\bigcup_i A_i \in \mathcal{G}$.

Thus $\mathcal{G}$ is a σ-algebra containing the open sets, hence it contains the Borel σ-algebra. Therefore $B \in \mathcal{G}$, so for the given $\varepsilon > 0$ there exists a compact $K \subseteq B$ with $\mu(B \setminus K) = \nu(B \setminus K) < \varepsilon$. This proves (i).

(ii) Outer regularity on Borel sets. Assume μ is finite on compact sets. Let $(K_m)_{m\geq 1}$ be the compact exhaustion from Lemma 1.6, so $K_m \subseteq \operatorname{int} K_{m+1}$ and $\bigcup_m K_m = X$. Set $U_m := \operatorname{int} K_m$.

Fix a Borel set $B \subseteq X$ and $\varepsilon > 0$. For each m the set $U_m \setminus B$ is Borel and has finite measure, because $\mu(U_m) \leq \mu(K_m) < \infty$ by local finiteness. Applying (i) to $U_m \setminus B$ yields a compact $C_m \subseteq U_m \setminus B$ with

$$\mu\Big((U_m \setminus B) \setminus C_m\Big) < \varepsilon/2^m.$$

Define

$$U := \bigcup_{m\geq 1} (U_m \setminus C_m),$$

which is open because each U_m is open and each C_m is closed. Since for every m we have $C_m \subseteq U_m \setminus B$, we have $U_m \cap B \subseteq U_m \setminus C_m$, and it follows that $B \subseteq U$ by taking union on both sides. Finally

$$\mu(U \setminus B) = \mu\Big(\bigcup_{m\geq 1}\Big((U_m \setminus C_m) \setminus B\Big)\Big) \leq \sum_{m\geq 1} \mu\Big((U_m \setminus B) \setminus C_m\Big) < \sum_{m\geq 1} \frac{\varepsilon}{2^m} = \varepsilon,$$

so U is the required open neighbourhood of B. This proves (ii). $\square$

Proposition 1.8 (A sufficient condition for Radon)**.** *Let X be a locally compact Hausdorff space in which every open set is σ-compact, and let μ^* be a Borel regular outer measure on X that is locally finite. Then:*

(i) μ^* is outer regular on all subsets of X;

(ii) μ^* is inner regular on all μ^* measurable sets.

In particular, μ^* is Radon.

Proof. Let $E \subseteq X$. By Borel regularity, $\exists$ Borel B_1 s.t. $E \subseteq B_1$ and $\mu^*(E) = \mu(B_1)$ By Lemma 1.7,

$$\begin{aligned} \mu^*(E) = \mu(B_1) &= \inf\{\mu^*(U) : B_1 \subseteq U,\ U \text{ open in } X\} \\ &\geq \inf\{\mu^*(U) : E \subseteq U,\ U \text{ open in } X\} \end{aligned}$$

The reverse inequality is immediate, proving outer regularity of E.

Assume now E is measurable. We first show inner regularity when $\mu(X) < \infty$. By Borel regularity, $\exists$ Borel B_2 s.t. $E^c \subseteq B_2$ and $\mu(E^c) = \mu(B_2)$. By finiteness of $\mu(X)$, it follows that $\mu(E) = \mu(B_2^c)$. Thus, noting that $B_2^c \subseteq E$, again by Lemma 1.7,

$$\begin{aligned} \mu(E) = \mu(B_2^c) &= \sup\{\mu^*(K) : K \subseteq B_2^c,\ K \text{ compact in } X\} \\ &\leq \sup\{\mu^*(K) : K \subseteq E,\ K \text{ compact in } X\}, \end{aligned}$$

proving inner regularity of E when $\mu(X) < \infty$

If $\mu(X) = \infty$, choose a compact exhaustion $(K_m)_{m\geq 1}$ with $X = \bigcup_m K_m$. Each $E \cap K_m$ has finite measure, so by the finite case $\mu(E \cap K_m)$ is the supremum of measures of compacts in $E \cap K_m$. Taking the increasing limit as $m \to \infty$ yields the inner regularity for E itself. □

Finally, we end this section with a useful condition to check if an outer measure is Borel in a metric space, i.e. if Borel sets are measurable.

Proposition 1.9 (Carathéodory's Criterion for Metric Spaces)**.** *Let (X, d) be a metric space and let μ^* be an outer measure on X. Suppose that for all $E, F \subseteq X$ with* $\operatorname{dist}(E, F) := \inf\{d(a, b) : a \in E, b \in F\} > 0$, *we have*

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F).$$

Then μ^* is a Borel outer measure.

Proof. We'll prove that all closed sets are measurable. Let $S \subseteq X$ and let $F \subseteq X$ be closed. We must show

$$\mu^*(S) \geq \mu^*(S \cap F) + \mu^*(S \cap F^c),$$

to conclude measurability of F

If $\mu^*(S) = \infty$, the inequality is trivial. Assume $\mu^*(S) < \infty$. For each $n \geq 1$, define the closed "thickening"

$$F_n := \{x \in X : d(x, F) \leq 1/n\}.$$

Then $\operatorname{dist}(S \cap F_n^c, S \cap F) > 1/n > 0$, so by hypothesis

$$\mu^*(S \cap F_n^c) + \mu^*(S \cap F) = \mu^*((S \cap F_n^c) \cup (S \cap F)) \leq \mu^*(S). \tag{1}$$

For $k \geq 1$, define

$$G_0 = \{x \in S : d(x, F) > 1\} \qquad G_k := \{x \in S : 1/(k+1) < d(x, F) \leq 1/k\}.$$

Let $x \in S \cap F^c$. As F is closed, $d(x, F) > 0$. If $x \notin F_n^c$, then $0 < d(x, F) \leq \frac{1}{n}$, which implies $x \in G_k$ for some $k \geq n$. Thus

$$S \cap F^c = (S \cap F_n^c) \cup \bigcup_{k \geq n} G_k. \tag{2}$$

Since $\operatorname{dist}(G_i, G_j) > 0$ for $|i - j| \geq 2$ by definition, the hypothesis gives:

$$\sum_{i=1}^{m} \mu^*\Big(G_{2i}\Big) = \mu^*\Big(\bigcup_{i=1}^{m} G_{2i}\Big) \leq \mu^*(S), \qquad \sum_{i=0}^{m} \mu^*\Big(G_{2i+1}\Big) = \mu^*\Big(\bigcup_{i=0}^{m} G_{2i+1}\Big) \leq \mu^*(S)$$

Therefore, $\sum_{k=1}^{\infty} \mu^*(G_k) \leq 2\mu^*(S) < \infty$. Thus, taking $n \to \infty$ in equation 2, we get

$$\lim_{n \to \infty} \mu^*(S \cap F_n^c) = \mu^*(S \cap F^c).$$

Finally, taking $n \to \infty$ in Equation 1 yields

$$\mu^*(S \cap F^c) + \mu^*(S \cap F) \leq \mu^*(S),$$

so F is μ-measurable. All closed sets are thus measurable, and thus all Borel sets are measurable. □

1.2 Covering Theorems

Covering theorems play a central role in GMT, allowing one to control sets by well-chosen families of balls or other simple sets. The most fundamental result in this direction is Vitali's covering theorem.

Theorem 1.10 (Vitali Covering Theorem)**.** *Let (X, d) be a separable metric space, and let $\mathcal{F}$ be a collection of balls (open or closed) such that*

$$M := \sup\{\operatorname{diam}(B) : B \in \mathcal{F}\} < \infty.$$

Then there exists a countable subcollection $\mathbf{G} \subseteq \mathcal{F}$ of pairwise disjoint balls such that

$$\bigcup_{B\in\mathcal{F}} B \;\subseteq\; \bigcup_{C\in\mathbf{G}} \tilde{C},$$

where $\tilde{C}$ denotes the ball concentric with C whose radius is 5 times that of C.

Proof. The proof follows a greedy selection procedure.

Step 1. Inductive construction. For $i \geq 1$, set

$$\mathcal{F}_i := \{B \in \mathcal{F} : M/2^i < \operatorname{diam}(B) \leq M/2^{i-1}\}.$$

From $\mathcal{F}_1$, choose a maximal disjoint subfamily $\mathbf{G}_1$. Having chosen $\mathbf{G}_1, \dots, \mathbf{G}_{k-1}$, pick $\mathbf{G}_k \subseteq \mathcal{F}_k$ to be maximal among families disjoint both internally and from $\mathbf{G}_1 \cup \cdots \cup \mathbf{G}_{k-1}$. Define

$$\mathbf{G} := \bigcup_{i=1}^{\infty} \mathbf{G}_i.$$

Step 2. Countability. By separability, for any $r > 0$, X can be covered by countably many balls of radius r. Since $\mathbf{G}_i$ consists of disjoint balls of diameter at least $M/2^i$, each $\mathbf{G}_i$ is at most countable. Hence $\mathbf{G}$ is countable.

Step 3. Covering property. Take $B \in \mathcal{F}$ not in $\mathbf{G}$. As $\{\mathcal{F}_k\}_{k\in\mathbb{N}}$ cover $\mathcal{F}$, $B \in \mathcal{F}_k$ for some k. By maximality of $\mathbf{G}_k$, the ball B must intersect some $C \in \mathbf{G}_j$ with $j \leq k$. By construction,

$$\operatorname{diam}(C) \;\geq\; M/2^k, \qquad \operatorname{diam}(B) \;\leq\; M/2^{k-1}.$$

Our goal is to enlarge C by a suitable factor so that it contains B. In the worst case, where $\operatorname{diam}(C) = M/2^k$ and $\operatorname{diam}(B) = M/2^{k-1}$, it is easy to see geometrically that enlarging C by a factor of 5 ensures $B \subseteq \tilde{C}$.

More precisely, write $C = B(c, r)$ and $B = B(b, \rho)$ so that $2r = \operatorname{diam}(C)$ and $2\rho = \operatorname{diam}(B)$. From $\operatorname{diam}(B) \leq 2\operatorname{diam}(C)$ we get $\rho \leq 2r$. Since B and C intersect, the distance between their centers satisfies

$$d(c, b) \leq r + \rho \leq r + 2r = 3r.$$

Take any point $y \in B$. Then

$$d(c, y) \leq d(c, b) + d(b, y) \leq 3r + \rho \leq 3r + 2r = 5r.$$

Hence every $y \in B$ lies in the concentric ball $B(c, 5r)$, i.e.

$$B \subseteq B(c, 5r) =: \widetilde{C}.$$

Thus every $B \in \mathcal{F}$ is contained in some $\tilde{C}$ with $C \in \mathbf{G}$, which proves the theorem. $\square$

Theorem 1.11. *Let (X,d) be a separable metric space and let ν be a Borel outer measure on X (i.e. Borel sets are measurable) which satisfies the ball-doubling property: there exists $C \geq 1$ such that for every $x \in X$ and every $r > 0$,*

$$\nu\Big(B(x,2r)\Big) \leq C\,\nu\Big(B(x,r)\Big).$$

Fix $\delta > 0$ and let $U \subseteq X$ be open. Then there exists a countable collection of pairwise disjoint balls $\{B_i\}_{i=1}^{\infty}$ with $\operatorname{diam}(B_i) < \delta$ and $B_i \subseteq U$ for all i, such that

$$\nu\Big(U \setminus \bigcup_{i=1}^{\infty} B_i\Big) = 0.$$

Proof. Assume first that $\nu(U) < \infty$. Choose for each $x \in U$ a radius $r_x > 0$ with

$$r_x < \min\{\tfrac{1}{2}\delta, \tfrac{1}{2}\operatorname{dist}(x, X \setminus U)\},$$

so that $B(x,r_x) \subseteq U$ and $\operatorname{diam} B(x,r_x) < \delta$. Let $\mathcal{F} := \{B(x,r_x) : x \in U\}$, which covers U.
By the doubling hypothesis,

$$\nu(\widetilde{B}) = \nu(B(x,5r)) \leq \nu(B(x,8r)) \leq C^3\,\nu(B(x,r)).$$

Set $A := C^3$. Applying Vitali' Covering Theorem 1.10 to the cover $\mathcal{F}$, we obtain a countable disjoint subcollection $\mathcal{G}_1 \subseteq \mathcal{F}$ such that

$$U \subseteq \bigcup_{B \in \mathcal{G}_1} \widetilde{B}.$$

Using disjointness and the inequality above,

$$\nu(U) \leq \sum_{B \in \mathcal{G}_1} \nu(\widetilde{B}) \leq A \sum_{B \in \mathcal{G}_1} \nu(B) = A\,\nu\Big(\bigcup_{B \in \mathcal{G}_1} B\Big).$$

Hence

$$\nu\Big(U \setminus \bigcup_{B \in \mathcal{G}_1} B\Big) = \nu(U) - \nu\Big(\bigcup_{B \in \mathcal{G}_1} B\Big) \leq \nu(U) - \frac{1}{A}\nu(U) = \Big(1 - \frac{1}{A}\Big)\nu(U).$$

Now iterate. Given the open residual set U_{n-1} with $\nu(U_{n-1}) < \infty$ (start with $U_0 := U$), choose the family of balls

$$\mathcal{F}^{(n)} := \{B(x,r_x) \in \mathcal{F} :\ B(x,r_x) \subseteq U_{n-1}\},$$

apply Vitali to obtain a countable disjoint subcollection $\mathcal{G}_n \subseteq \mathcal{F}^{(n)}$ with

$$U_{n-1} \subseteq \bigcup_{B\in\mathcal{G}_n} \widetilde{B},$$

and set $U_n := U_{n-1} \setminus (\bigcup_{B\in\mathcal{G}_n} B)$. The same estimate as above yields

$$\nu(U_n) \leq \Big(1 - \frac{1}{A}\Big)\nu(U_{n-1}).$$

Iterating,

$$\nu(U_n) \leq \Big(1 - \frac{1}{A}\Big)^n \nu(U_0) \xrightarrow[n\to\infty]{} 0,$$

since $0 < 1 - \frac{1}{A} < 1$.

Finally let $\{B_i\} := \bigcup_{n\geq 1} \mathcal{G}_n$. This is a countable union of countable disjoint families, hence countable, the balls are pairwise disjoint, each lies in U and has diameter $< \delta$, and

$$\nu\Big(U \setminus \bigcup_i B_i\Big) = \lim_{n\to\infty} \nu(U_n) = 0.$$

This completes the proof for $\nu(U) < \infty$. If $\nu(U) = \infty$, fix $x_0 \in U$, and apply the above construction to each set in the countable collection $U_r = \{x \in U : r-1 \leq d(x, x_0) < r\}$ for $r \in \mathbb{N}$. $\square$

Corollary 1.12. *Let $U \subseteq \mathbb{R}^n$ be open and fix $\delta > 0$. Then there exists a countable collection of pairwise disjoint balls $\{B_i\}_{i=1}^\infty$ with $\operatorname{diam}(B_i) < \delta$ and $B_i \subseteq U$ for all i, such that*

$$\mathcal{L}^n\Big(U \setminus \bigcup_{i=1}^\infty B_i\Big) = 0,$$

where $\mathcal{L}^n$ denotes Lebesgue measure on $\mathbb{R}^n$.

Although Corollary 1.12 will be sufficient for our purposes in establishing the area formula, it is worth mentioning another powerful covering result in $\mathbb{R}^n$, namely the *Besicovitch covering theorem*. This theorem may be useful for future reports.

The main limitation of Vitali's covering theorem is the need to enlarge the selected balls by a fixed factor (e.g., 5), which restricts its applicability to measures that satisfy the ball-doubling property from Theorem 1.11. By contrast, the Besicovitch covering theorem does not need to enlarge the ball, and provides an analogue of Theorem 1.11 in $\mathbb{R}^n$ that holds for *any* Borel outer measure, without assuming doubling. Its proof is substantially more involved than Vitali's and doesn't really provide any additional insight into the statement. While it follows the same general "greedy selection" scheme, it makes essential use of the geometry peculiar to $\mathbb{R}^n$. Since we will not need this result in proving the area formula, we only sketch its ideas rather than present the full argument.

Theorem 1.13 (Besicovitch Covering Theorem)**.** *Let $E \subseteq \mathbb{R}^n$ be a bounded set, and let $\mathcal{F}$ be a family of open/closed balls $\{B(x, r_x) : x \in E\}$ with centers in E and uniformly* *bounded radii* $\sup_{x\in E} r_x < \infty$. *Then there exists a constant* $N = N(n)$, *depending only on the dimension* n *(and not the cover), and subfamilies*

$$\mathcal{F}_1, \dots, \mathcal{F}_N \subseteq \mathcal{F}$$

such that:

1. *Each* $\mathcal{F}_i$ *consists of pairwise disjoint balls;*
2. $E \subseteq \bigcup_{i=1}^{N} \bigcup_{B\in\mathcal{F}_i} B$.

Sketch of proof. We give only a rough outline of the argument.

Step 1 Construction: Assume E is bounded and set

$$D := \sup\{\mathrm{diam}(B) : B \in \mathcal{F}\}.$$

Choose a first ball $B_1 \in \mathcal{F}$ whose diameter is at least $\frac{3}{4}D$. Inductively, suppose $B_1, \dots, B_{j-1}$ have been chosen. Define

$$E_j := E \setminus \bigcup_{i=1}^{j-1} B_i.$$

If $E_j \neq \varnothing$, pick a ball $B_j \in \mathcal{F}$ centered in E_j such that

$$\mathrm{diam}(B_j) \geq \tfrac{3}{4} \sup\{\mathrm{diam}(B) : B \in \mathcal{F},\ \mathrm{center}(B) \in E_j\}.$$

In words, each new ball has nearly maximal diameter among those whose centers still lie outside the previously chosen ones. This produces a sequence $\{B_j\}$, and by construction one can check that

$$E \subseteq \bigcup_{j=1}^{\infty} B_j.$$

Step 2 Geometry of overlaps: The selected family $\{B_j\}$ need not be disjoint. However, the key fact—specific to the geometry of $\mathbb{R}^n$—is that for each k, the set of indices

$$I_k := \{j \leq k : B_j \cap B_k \neq \varnothing\}$$

has uniformly bounded cardinality: $|I_k| \leq N(n)$, where $N(n)$ depends only on the dimension. Establishing this uniform bound requires a rather lengthy and ad-hoc argument, and forms the key step of the proof.

Step 3 Decomposition into disjoint subfamilies: From the uniform overlap bound in Step 2, it follows that we can partition the family $\{B_j\}$ into $N(n)$ disjoint subcollections $\mathcal{F}_1, \dots, \mathcal{F}_{N(n)}$, each consisting of disjoint balls. Together, these subcollections cover E, as required. □

An important consequence is analogue of Theorem 1.11 in $\mathbb{R}^n$ for *any* Borel outer measure, without assuming doubling.

Theorem 1.14. *Let ν be a Borel outer measure on $\mathbb{R}^n$, and let $\mathcal{F}$ be a collection of balls in $\mathbb{R}^n$. Denote by E the set of centers of balls in $\mathcal{F}$. Suppose that $\nu(E) < \infty$, and*

$$\inf\{r > 0 : B(x,r) \in \mathcal{F}\} = 0 \quad \textit{for each } x \in E \tag{1}$$

Then for every open set $U \subseteq \mathbb{R}^n$ there exists a countable subcollection $\mathcal{G} \subseteq \mathcal{F}$ of pairwise disjoint balls such that

$$\bigcup_{B\in\mathcal{G}} B \subseteq U \quad \textit{and} \quad \nu\Big((E\cap U) \setminus \bigcup_{B\in\mathcal{G}} B\Big) = 0.$$

Proof. The construction follows the same pattern as in Theorem 1.11, but now we appeal to Besicovitch's covering theorem in place of the Vitali's covering theorem..
1. Consider the subfamily

$$\mathcal{F}_1 := \{B \in \mathcal{F} : \mathrm{diam}(B) \leq 1,\ B \subseteq U\}.$$

By the hypothesis 1 and as U is open, $\mathcal{F}_1$ covers $E \cap U$. Now, by Besicovitch's theorem, there exist finitely many disjoint families of balls $\mathcal{G}_1, \dots, \mathcal{G}_{N(n)} \subseteq \mathcal{F}_1$ such that

$$E \cap U \subseteq \bigcup_{i=1}^{N(n)} \bigcup_{B\in\mathcal{G}_i} B.$$

Hence

$$\nu(E\cap U) \leq \sum_{i=1}^{N(n)} \nu\Big(E \cap U \cap \bigcup_{B\in\mathcal{G}_i} B\Big).$$

Thus there exists some $j \in \{1, \dots, N(n)\}$ with

$$\nu\Big(E \cap U \cap \bigcup_{B\in\mathcal{G}_j} B\Big) \geq \tfrac{1}{N(n)} \nu(E\cap U).$$

2. From this family $\mathcal{G}_j$ choose finitely many balls $B_1, \dots, B_{M_1}$ so that

$$\nu\Big(E \cap U \cap \bigcup_{i=1}^{M_1} B_i\Big) \geq \Big(\frac{1}{2N(n)}\Big)\nu(E\cap U),$$

Now,

$$\nu(E\cap U) \ = \ \nu\left(E \cap U \cap \bigcup_{i=1}^{M_1} B_i\right) \ + \ \nu\left((E\cap U) \setminus \bigcup_{i=1}^{M_1} B_i\right),$$

since $\bigcup_{i=1}^{M_1} B_i$ is ν-measurable, as ν is a Borel outer measure. Thus:

$$\nu\Big((E\cap U) \setminus \bigcup_{i=1}^{M_1} B_i\Big) \leq \Big(1 - \frac{1}{2N(n)}\Big)\, \nu(E\cap U).$$

3. Define $U_2 := U \setminus \bigcup_{i=1}^{M_1} B_i$ and repeat the same construction with the family $\mathcal{F}_2 := \{B \in \mathcal{F} : \operatorname{diam}(B) \leq \delta,\ B \subseteq U_2\}$. This yields further disjoint balls $B_{M_1+1}, \ldots, B_{M_2}$ such that

$$\nu\Big((E \cap U) \setminus \bigcup_{i=1}^{M_2} B_i\Big) \leq \Big(1 - \frac{1}{2N(n)}\Big)\, \nu(E \cap U_2) \leq \Big(1 - \frac{1}{2N(n)}\Big)^2 \nu(E \cap U).$$

4. Iterating this argument produces a countable collection of pairwise disjoint balls $\{B_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ with $B_i \subseteq U$ and

$$\nu\Big((E \cap U) \setminus \bigcup_{i=1}^{\infty} B_i\Big) = 0,$$

since $\Big(1 - \frac{1}{2N(n)}\Big)^k \nu(E \cap U) \to 0$ as $k \to \infty$ and $\nu(E) < \infty$. This completes the proof. $\square$

2 Hausdorff Measures

2.1 Definition and Properties

The Hausdorff measure provides a way to assign a meaningful "s-dimensional volume" to subsets of a metric space. The basic idea is to cover a set by small pieces, measure each piece in terms of its diameter raised to the power s, and then optimize by taking the infimum over all such coverings. This procedure, as we shall see, does indeed generalize other notions of volume of integer dimension we might have encountered. The idea discussed is made precise in the following definition:

Definition 2.1 (s-dimensional Hausdorff measure)**.** Let (X, d) be a metric space, $A \subset X$, and $s \in \mathbb{R}$ s.t. $s \geq 0$. For $\delta > 0$ define

$$\mathcal{H}^s_\delta(A) := \inf \left\{ \alpha(s) \sum_{j=1}^{\infty} \left(\frac{\operatorname{diam} C_j}{2} \right)^s \ : \ A \subset \bigcup_{j=1}^{\infty} C_j, \ \operatorname{diam} C_j \leq \delta \right\},$$

where $\operatorname{diam} C := \sup\{d(x,y) : x, y \in C\}$, and the infimum is taken over all countable covers of A by sets of diameter at most δ.
The *s-dimensional Hausdorff measure* is then

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

The normalization constant is

$$\alpha(s) \ = \ \frac{\pi^{s/2}}{\Gamma\left(\frac{s}{2} + 1\right)}.$$

where Γ denotes the Gamma function.

Remark 2.2. For integer n, the constant $\alpha(n)$ is exactly the volume of the unit ball in $\mathbb{R}^n$. As we shall see, we introduce this normalization so that the n-dimensional Hausdorff measure $\mathcal{H}^n$ coincides with Lebesgue measure $\mathcal{L}^n$ on $\mathbb{R}^n$, rather than being a scaled version of it. A main goal of this chapter will be to establish this equality $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$.

Proposition 2.3. *Let (X, d) be a metric space and $s \geq 0$. Then $\mathcal{H}^s$ is an outer measure on X.*

Proof. It is immediate from the definition that $\mathcal{H}^s_\delta(\varnothing) = 0 \ \forall \delta > 0 \implies \mathcal{H}^s(\varnothing) = 0$.

If $A \subset B$, any cover of B is also a cover of A, so $\mathcal{H}^s_\delta(A) \leq \mathcal{H}^s_\delta(B)$. Taking supremum over $\delta > 0$, we get $\mathcal{H}^s(A) \leq \mathcal{H}^s(B)$.

Let $A = \bigcup_{j=1}^{\infty} A_j$ and fix $\varepsilon > 0$. For each j, choose δ-covers $\{C_{j,k}\}_{k \geq 1}$ of A_j such that

$$\sum_{k=1}^{\infty} \alpha(s) \left(\frac{\operatorname{diam} C_{j,k}}{2} \right)^s \leq \mathcal{H}^s_\delta(A_j) + \frac{\varepsilon}{2^j}.$$

Then $\bigcup_{j,k} C_{j,k}$ is a δ-cover of A, so

$$\mathcal{H}^s_\delta(A) \leq \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \alpha(s)\left(\frac{\operatorname{diam} C_{j,k}}{2}\right)^s \leq \sum_{j=1}^{\infty} \mathcal{H}^s_\delta(A_j) + \varepsilon.$$

Taking $\delta \downarrow 0$ and then $\varepsilon \to 0$, we obtain

$$\mathcal{H}^s(A) \leq \sum_{j=1}^{\infty} \mathcal{H}^s(A_j),$$

which proves countable subadditivity. □

Remark 2.4. Although $\mathcal{H}^n$ is strictly speaking an outer measure on X, we shall simply refer to it as the *Hausdorff measure* rather than the *Hausdorff outer measure*. The Borel sets are indeed $\mathcal{H}^n$-measurable; the next proposition shows that $\mathcal{H}^n$ is Borel regular.

Proposition 2.5. *Let (X, d) be a metric space and $0 \leq s < \infty$. Then $\mathcal{H}^s$ is a Borel regular outer measure.*

Proof. Step 1: Borel sets are measurable: We shall show this using Carathéodory's criterion (Proposition 1.9). Let $E, F \subset X$ with $\operatorname{dist}(E, F) > 0$. Fix δ with $0 < \delta < \operatorname{dist}(E, F)$.

Let $\{C_j\}$ be a countable cover of $E \cup F$ such that $\operatorname{diam} C_j \leq \delta$. Let $A_1 = \{j : C_j \cap E \neq \varnothing\}$ and $A_2 = \{j : C_j \cap F \neq \varnothing\}$. Note that $\{C_j\}_{j \in A_1}$ and$\{C_j\}_{j\in A_2}$ form covers of E and F resp. Moreover, if $j_1 \in A_1$, $j_2 \in A_2$, then $C_{j_1} \cap C_{j_2} = \varnothing$. Indeed, because if not, C_{j_1} would contain an $x \in E$ and a $y \in F$, giving $\operatorname{diam} C_j \geq d(x, y) \geq \operatorname{dist}(E, F) > \delta$, contradicting the requirement $\operatorname{diam} C_j \leq \delta$. Thus:

$$\sum_{j=1}^{\infty}\left(\frac{\operatorname{diam} C_j}{2}\right)^s = \sum_{C_j \cap E \neq \varnothing}\left(\frac{\operatorname{diam} C_j}{2}\right)^s + \sum_{C_j \cap F \neq \varnothing}\left(\frac{\operatorname{diam} C_j}{2}\right)^s \geq \mathcal{H}^s_\delta(E) + \mathcal{H}^s_\delta(F)$$

Taking infimum over all δ-covers of $E \cup F$ gives

$$\mathcal{H}^s_\delta(E \cup F) \geq \mathcal{H}^s_\delta(E) + \mathcal{H}^s_\delta(F).$$

The other inequality is immediate as $\mathcal{H}_\delta$ is an outer measure, and we get:

$$\mathcal{H}^s_\delta(E \cup F) = \mathcal{H}^s_\delta(E) + \mathcal{H}^s_\delta(F).$$

Taking the limit as $\delta \downarrow 0$, we obtain

$$\mathcal{H}^s(E \cup F) = \mathcal{H}^s(E) + \mathcal{H}^s(F).$$

By Carathéodory's criterion (Proposition 1.9), it follows that all Borel sets are $\mathcal{H}^s$-measurable.

Step 2: Borel regularity: Let $A \subset X$ and fix $n \geq 1$ an integer. Observe that $\operatorname{diam}(C) = \operatorname{diam}(\overline{C})$, and thus in the definition of $\mathcal{H}^s_\delta$, one can replace the arbitrary cover $\{C_j\}$ by a closed cover. Thus, $\exists$ countable cover of closed sets $\{C_j\}$ of A s.t.

$$\sum_j \alpha(s)\Big(\frac{\operatorname{diam} C_j}{2}\Big)^s \leq \mathcal{H}^s_\delta(A) + \frac{1}{n}.$$

Define

$$B_n := \bigcup_j C_j$$

which is a Borel set containing A and satisfying

$$\mathcal{H}^s_\delta(B_n) \leq \sum_j \alpha(s)\Big(\frac{\operatorname{diam} \overline{C_j}}{2}\Big)^s \leq \mathcal{H}^s_\delta(A) + \frac{1}{n}.$$

Now, let

$$B := \bigcap_{n=1}^{\infty} B_n.$$

Then B is a Borel set, $A \subset B$, and taking $\delta \downarrow 0$ and $n \to \infty$, we obtain

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

proving the existence of a Borel superset with the same Hausdorff measure. This establishes Borel regularity.

□

2.2 Hausdorff Dimension and Examples

Definition 2.6 (Hausdorff dimension)**.** Let $A \subset X$ be a subset of a metric space (X, d). The *Hausdorff dimension* of A is defined as

$$\dim_{\mathcal{H}}(A) := \inf\{s \geq 0 : \mathcal{H}^s(A) = 0\} = \sup\{s \geq 0 : \mathcal{H}^s(A) = \infty\}.$$

The second equality is justified by the proposition below.

Proposition 2.7 (Monotonicity of Hausdorff measure)**.** *Let $A \subset X$ and $s_0 \geq 0$. Then:*

(i) If $\mathcal{H}^{s_0}(A) > 0$, then $\forall s < s_0$, $\mathcal{H}^s(A) = \infty$.

(ii) If $\mathcal{H}^{s_0}(A) < \infty$, $\forall s > s_0$, $\mathcal{H}^s(A) = 0$.

Proof. (i) Suppose $\mathcal{H}^{s_0}(A) > 0$ and let $s < s_0$. Then for any δ-cover $\{C_j\}$ of A (i.e.

$(\mathrm{diam}(C_j) < \delta)$,

$$\begin{aligned}\sum_j \alpha(s)(\mathrm{diam}\, C_j/2)^s &= \sum_j \alpha(s_0)(\mathrm{diam}\, C_j/2)^{s_0} \cdot (\mathrm{diam}\, C_j/2)^{s-s_0} \\ &\geq \left(\frac{\delta}{2}\right)^{s-s_0} \sum_j \alpha(s_0)(\mathrm{diam}\, C_j/2)^{s_0}\end{aligned}$$

Taking infimum over all such δ-covers, we get:

$$\mathcal{H}^s_\delta(A) \geq \left(\frac{\delta}{2}\right)^{s-s_0} \mathcal{H}^{s_0}_\delta(A)$$

As $s - s_0 < 0$ and $\mathcal{H}^{s_0}(A) > 0$, letting $\delta \to 0$, we get $\mathcal{H}^s(A) = \infty$
The proof of (ii) is analogous. $\square$

Corollary 2.8. *Let $A \subset X$ and $s := \dim_{\mathcal{H}}(A)$. Then*

$$\mathcal{H}^t(A) = \begin{cases} \infty, & t < s, \\ 0, & t > s, \\ \textit{anything in } [0, \infty], & t = s. \end{cases}$$

Remark 2.9. Note in particular that $\mathcal{H}^s(A)$, where $s := \dim_{\mathcal{H}}(A)$, need not be finite! For example, as we shall see in the next proposition, $\mathcal{H}^1 = \mathcal{L}^1$ in $\mathbb{R}^1$ and $\mathcal{H}^s(A) = 0$ for $s > n$, $A \subset \mathbb{R}^n$. Thus taking $A = \mathbb{R}$, we can conclude that $\dim_{\mathcal{H}}(\mathbb{R}) = 1$ and $\mathcal{H}^1(\mathbb{R}) = \infty$.

Below are some simple properties which shall be useful later:

Proposition 2.10 (Basic Properties)**.**

(i) In a metric space X, $\mathcal{H}^0$ = counting measure

(ii) $\mathcal{H}^1 = \mathcal{L}^1$ in $\mathbb{R}$

Proof. (i) Let $x \in X$. Noting that $\alpha(0) = 1$, and taking cover $\{C_j\}$ with $C_1 = \{x\}$ and $C_j = \varnothing$ for $j > 1$, we get $\mathcal{H}^0(x) = 1$. By countable additivity, (i) follows.

(ii) Note that $\alpha(1) = \frac{\pi^{1/2}}{\Gamma(3/2)} = 2$, and thus the expression inside the summation of $\mathcal{H}^s_\delta$ is just $\mathrm{diam}(C_i)$. Fix $A \subset \mathbb{R}$ and an arbitrary $\delta > 0$. By definition of $\mathcal{L}^1$,

$$\begin{aligned}\mathcal{L}^1(A) &= \inf\left\{\sum_{j=1}^\infty \mathrm{diam}\, C_j : \; A \subset \bigcup_{j=1}^\infty C_j, \; C_j \text{ closed interval}\right\}, \\ &= \inf\left\{\sum_{j=1}^\infty \mathrm{diam}\, C_j : \; A \subset \bigcup_{j=1}^\infty C_j\right\}, \\ &\leq \inf\left\{\sum_{j=1}^\infty \mathrm{diam}\, C_j : \; A \subset \bigcup_{j=1}^\infty C_j, \, \mathrm{diam}(C_j) < \delta\right\}, \\ &= \mathcal{H}^1_\delta(A)\end{aligned}$$

where the equality in the second line follows as any set C_j with $\mathrm{diam}(C_j) < \delta$ is contained in an closed interval with the same diameter.

Since this holds for every $\delta > 0$, we get $\mathcal{L}^1(A) \leq \mathcal{H}^1(A)$.
It remains to show the other inequality. Let $\{C_j\}_{j\geq 1}$ be any countable cover of A by closed intervals. Partition $\mathbb{R}$ into the intervals

$$I_k := [k\delta, (k+1)\delta] \qquad (k \in \mathbb{Z}).$$

For each pair (j,k) set $C_{j,k} := C_j \cap I_k$. Each nonempty $C_{j,k}$ is a closed interval with $\mathrm{diam}(C_{j,k}) \leq \delta$, and for fixed j the family $\{C_{j,k}\}_k$ is pairwise disjoint and covers C_j. Hence

$$\sum_{k\in\mathbb{Z}} \mathrm{diam}(C_{j,k}) = \mathrm{diam}(C_j).$$

The refined family $\{C_{j,k} : j \geq 1,\ k \in \mathbb{Z},\ C_{j,k} \neq \varnothing\}$ is a cover of A by sets of diameter $\leq \delta$. Therefore, for the given cover $\{C_j\}$ we have

$$\sum_{j\geq 1} \mathrm{diam}(C_j) = \sum_{j,k} \mathrm{diam}(C_{j,k}) \geq \mathcal{H}^1_\delta(A).$$

Taking the infimum over all covers $\{C_j\}$ of A where C_j are closed intervals yields

$$\mathcal{L}^1(A) \geq \mathcal{H}^1_\delta(A).$$

Passing to the limit $\delta \downarrow 0$ gives $\mathcal{L}^1(A) \geq \mathcal{H}^1(A)$, as required.

□

Having explored some basic properties of Hausdorff measures, we now conclude this section by looking at examples of sets and their corresponding Hausdorff measures. As we shall see later, Theorem 3.22 shows that any C^1 k-submanifold of $\mathbb{R}^n$ has Hausdorff dimension k, providing a broad class of examples with integer Hausdorff dimension. In contrast, there also exist sets whose Hausdorff dimension is non-integer. We present a general construction to obtain some of these examples below (from [1], [5]). Before that, we prove some preliminary results.

Definition 2.11 (Hausdorff distance)**.** Let (M,d) be a metric space and let $K(M)$ denote the family of non-empty compact subsets of M. The *Hausdorff distance* between $A, B \in K(M)$ is defined by

$$d_H(A,B) := \max\left\{ \sup_{a\in A} \inf_{b\in B} d(a,b),\ \sup_{b\in B} \inf_{a\in A} d(a,b) \right\}.$$

It is straightforward to check that d_H defines a metric on $K(M)$.

Lemma 2.12. *Let $f : M \to M$ be a contraction, i.e. f is Lipschitz with $0 < c < 1$ where $c := \mathrm{Lip}(f)$ Then the induced map*

$$f : K(M) \to K(M), \qquad X \mapsto f(X),$$

is also a contraction with the same constant c, i.e.

$$d_H(f(A), f(B)) \leq c\, d_H(A, B) \qquad \forall\, A, B \in K(M).$$

Proof. Let $a \in A$. Then

$$\inf_{b \in B} d(f(a), f(b)) \leq \inf_{b \in B} c\, d(a, b) = c \inf_{b \in B} d(a, b).$$

Taking the supremum over $a \in A$ yields

$$\sup_{a \in A} \inf_{b \in B} d(f(a), f(b)) \leq c \sup_{a \in A} \inf_{b \in B} d(a, b).$$

The other term in d_H is handled similarly, giving the claim. $\square$

Proposition 2.13. *Let $\{f_i\}_{i=1}^k$ be a finite family of contractions on M with contraction constants $c_i < 1$. Set $c = \max_i c_i < 1$ and define*

$$F : K(M) \to K(M), \qquad F(X) := \bigcup_{i=1}^{k} f_i(X).$$

Then F is a contraction on $(K(M), d_H)$ with constant c.

Proof. Let $A, B \in K(M)$ Set

$$S_1 := \sup_{x \in F(A)} \inf_{y \in F(B)} d(x, y), \qquad S_2 := \sup_{y \in F(B)} \inf_{x \in F(A)} d(x, y).$$

By compactness of $F(A)$ and $F(B)$ each supremum is attained. Thus there exists $x_0 \in F(A)$ with

$$S_1 = \inf_{y \in F(B)} d(x_0, y).$$

By definition of $F(A)$, $x_0 \in f_{j_0}(A)$ for some index j_0. Hence

$$S_1 = \inf_{y \in F(B)} d(x_0, y) \leq \inf_{y \in f_{j_0}(B)} d(x_0, y) \leq \sup_{x \in f_{j_0}(A)} \inf_{y \in f_{j_0}(B)} d(x, y).$$

Since the last bound holds for some j_0, we obtain

$$S_1 \leq \max_{1 \leq i \leq k} \sup_{x \in f_i(A)} \inf_{y \in f_i(B)} d(x, y).$$

Repeating the same reasoning for S_2 gives

$$S_2 \le \max_{1\le i\le k} \sup_{y\in f_i(B)} \inf_{x\in f_i(A)} d(x,y).$$

Taking the maximum of the two inequalities yields

$$d_H(F(A),F(B)) = \max\{S_1,S_2\} \le \max_{1\le i\le k} d_H(f_i(A),f_i(B)).$$

Combining this with Lemma 2.12 yields

$$d_H(F(A),F(B)) \le c\, d_H(A,B),$$

as required. □

Definition 2.14. In $\mathbb{R}^n$, a *similitude* is a map of the form

$$\Psi(x) = \lambda R(x) + a,$$

where $\lambda > 0$ is the scaling factor, $R \in O(n)$ is an orthogonal transformation, and $a \in \mathbb{R}^n$ is a translation vector.

Lemma 2.15. *Let $\Psi(x) = \lambda R(x) + a$ be a similitude of $\mathbb{R}^n$, with $\lambda > 0$, $R \in O(n)$ and $a \in \mathbb{R}^n$. Then for every $s \ge 0$ and every $A \subset \mathbb{R}^n$,*

$$\mathcal{H}^s\Big(\Psi(A)\Big) = \lambda^s\, \mathcal{H}^s(A).$$

Proof. Since R is orthogonal and translation by a preserves distances, we have for all x, y

$$d\Big(\Psi(x),\Psi(y)\Big) = \lambda\, d(x,y).$$

Fix $\delta > 0$. If $\{U_j\}$ is any cover of A with $\operatorname{diam} U_j \le \delta$ then $\{\Psi(U_j)\}$ is a cover of $\Psi(A)$ with $\operatorname{diam} \Psi(U_j) = \lambda \operatorname{diam} U_j \le \lambda\delta$. Hence

$$\mathcal{H}^s_{\lambda\delta}\Big(\Psi(A)\Big) \le \sum_j (\operatorname{diam} \Psi(U_j))^s = \lambda^s \sum_j (\operatorname{diam} U_j)^s.$$

Taking the infimum over such covers and then letting $\delta \downarrow 0$ gives

$$\mathcal{H}^s\Big(\Psi(A)\Big) \le \lambda^s \mathcal{H}^s(A).$$

Applying the same argument to Ψ^{-1} (which is a similitude with scaling factor $1/\lambda$) yields the reverse inequality, and equality follows. □

With these preliminaries out of the way, we can now define a self-similar fractal and prove its existence.

Definition 2.16 (Self-similar fractal)**.** Let $\{\Psi_i\}_{i=1}^k$ be a finite collection of similitudes of $\mathbb{R}^n$ with scaling factors $\lambda_i \in (0,1)$. The *self-similar fractal* (or set) associated with $\{\Psi_i\}$ is the unique compact set C satisfying

$$C = \bigcup_{i=1}^{k} \Psi_i(C).$$

Proof of Existence of C. The existence and uniqueness of such C follow immediately from Proposition 2.13 and Banach's fixed point theorem applied to the map $X \mapsto \bigcup_i \Psi_i(X)$. □

Proposition 2.17. *Let C be the self-similar fractal associated with $\{\Psi_i\}_{i=1}^k$. Assume that $\{\Psi_i(C)\}_{i=1}^k$ are pairwise disjoint and $0 < \mathcal{H}^d(C) < \infty$ where $d = \dim_{\mathcal{H}}(C)$. Then d is the unique real number s.t.*

$$\sum_{i=1}^{k} \lambda_i^d = 1.$$

Proof. The proof is straightforward. Let $\Psi_i = \lambda_i R + a_i$. Then, by Lemma 2.15, $\mathcal{H}^d(\Psi_i(C)) = \lambda_i^d \mathcal{H}^n(C)$.

Now, as C, $\Psi_i(C)$ are compact, they are $\mathcal{H}^n$ measurable. Thus, by countable additivity,

$$\mathcal{H}^d(C) = \sum_{i=1}^{k} \mathcal{H}^d(C_i) = \sum_{i=1}^{k} \lambda_i^d \mathcal{H}^d(C_i)$$

As $0 < \mathcal{H}^d(C_i) < \infty$, we get the desired result. □

In practice, the hardest part in computing the dimension of a self-similar fractal is establishing that $0 < \mathcal{H}^d(C) < \infty$, as we see in the next example.

Example 2.18 (Cantor set)**.** We use the standard level-n construction of C. Let $C_0 = [0,1]$, and recursively define C_{n+1} by removing the open middle third of each interval in C_n. Then $C = \bigcap_{n=0}^{\infty} C_n$, and C_n consists of 2^n closed intervals of length 3^{-n}. Defining the two similitudes

$$\Psi_1(x) = \tfrac{1}{3}x, \qquad \Psi_2(x) = \tfrac{1}{3}x + \tfrac{2}{3},$$

we see that $C = \Psi_1(C) \cup \Psi_2(C)$, and the pieces $\Psi_1(C)$ and $\Psi_2(C)$ are disjoint. Hence, assuming $0 < \mathcal{H}^{dim_{\mathcal{H}} C}(C) < \infty$, the hypothesis of Proposition 2.17 holds. Then Hausdorff dimension d of C is the unique solution of

$$2\left(\tfrac{1}{3}\right)^d = 1,$$

i.e.

$$d = \frac{\log 2}{\log 3} \approx 0.63093.$$

It remains to show $0 < \mathcal{H}^d(C) < \infty$.

For any $n \geq 0$, the intervals in C_n form a cover of C by sets of diameter 3^{-n}. Thus,

$$\mathcal{H}^d_{3^{-n}}(C) \leq \alpha(d) \sum_{I \in C_n} (\operatorname{diam} I)^d = \alpha(d)\, 2^n (3^{-n})^d = \alpha(d)\, (2 \cdot 3^{-d})^n = \alpha(d),$$

since $2 \cdot 3^{-d} = 1$ by definition of d. Letting $n \to \infty$ gives

$$\mathcal{H}^d(C) = \lim_{\delta \to 0} \mathcal{H}^d_\delta(C) \leq \alpha(d) < \infty.$$

For the converse, first note that we can replace arbitrary covers by open/closed covers in the definition of Hausdorff measure. Let $\{U_i\}_{i \in I}$ be any open cover of C with $\operatorname{diam}(U_i) < \delta$. By compactness of C, we can extract a finite subcover $\{U_1, \ldots, U_m\}$. Set

$$\delta' := \min_{1 \leq j \leq m} \operatorname{diam}(U_j) > 0.$$

Choose n large enough s.t. $3^{-n} < \delta'$. Then each $J_k \in C_n$ must be contained in some U_j. Since the intervals J_k are disjoint, we get:

$$\sum_{j=1}^{m} (\operatorname{diam} U_j)^d \geq \sum_{k=1}^{2^n} (\operatorname{diam} J_k)^d = 2^n (3^{-n})^d = 1.$$

Taking the infimum over all open covers and letting $n \to \infty$, we conclude that

$$\mathcal{H}^d_\delta(C) \geq \alpha(d) > 0.$$

for each δ, and thus $\mathcal{H}^d(C) \geq \alpha(d)$
Thus, we can conclude that

$$\mathcal{H}^d(C) = \alpha(d) \text{ and } d = \frac{\log 2}{\log 3},$$

2.3 $\mathcal{H}^n = \mathcal{L}^n$

We conclude this chapter by establishing the equivalence we set out to show, namely that $\mathcal{H}^n = \mathcal{L}^n$. Recall that, by the definition of the product outer measure,

$$\begin{aligned}
\mathcal{L}^n(E) &= \inf \left\{ \sum_{j=1}^{\infty} |R_j| : E \subset \bigcup_{j=1}^{\infty} R_j,\ R_j \text{ rectangle} \right\} \\
&= \inf \left\{ \sum_{j=1}^{\infty} |Q_j| : E \subset \bigcup_{j=1}^{\infty} Q_j,\ Q_j \text{ cube} \right\} \\
&= \inf \left\{ \sum_{j=1}^{\infty} |Q_j| : E \subset \bigcup_{j=1}^{\infty} Q_j,\ Q_j \text{ cube},\ \operatorname{diam}(Q_j) < \delta \right\},
\end{aligned}$$

$$\mathcal{L}^n(E) = \inf\Big\{\sum_{j=1}^{\infty} |Q_j| : E \subset \bigcup_{j=1}^{\infty} Q_j,\ Q_j \text{ cube},\ \operatorname{diam}(Q_j) < \delta\Big\},$$

where $|R_j|, |Q_j|$ denote. The second equality follows as any rectangle R can be covered by finitely many cubes with total volume arbitrarily close $|R|$. The third follows in the same way we showed for the case $n = 1$ when proving $\mathcal{H}^1 = \mathcal{L}^1$ in Proposition 2.10: Indeed, given a covering of E by rectangles, we may partition $\mathbb{R}^n$ into a grid of disjoint cubes of diameter less than δ, and then refine each covering rectangle into a union of such cubes. In this way we obtain a refinement of the original covering of E by cubes with diam $< \delta$, without increasing the total volume of the covering family.

The first step in the proof is to establish the isodiametric inequality:

Theorem 2.19 (Isodiametric Inequality). *For every set $E \subset \mathbb{R}^n$,*

$$\mathcal{L}^n(E) \leq \alpha(n)\Big(\tfrac{\operatorname{diam}(E)}{2}\Big)^n = \operatorname{vol}(B(0,R))$$

where $R = \frac{\text{diam(E)}}{2}$

Remark 2.20. The key insight we used in the proof that $\mathcal{H}^1 = \mathcal{L}^1$ in $\mathbb{R}^1$ (Proposition 2.10), is the fact than an arbitrary set C with $\operatorname{diam}(C) < \delta$ is contained in an interval with the same diameter. This observation makes the isodiametric immediate when $n = 1$. However, the fact that a set $C \subset \mathbb{R}^n$ of $\operatorname{diam}(C) < \delta$ can be contained in a ball with the same diameter is far from true in $\mathbb{R}^n$ with $n > 1$.

To address this, consider a more specific class of sets E, namely those which are symmetric about a point, say the origin. Then

$$E \subset B(0, \frac{\text{diam(E)}}{2}) \tag{1}$$

Indeed, if $x \in E \setminus B(0, \frac{\text{diam(E)}}{2})$, then by symmetry of E, $-x \in E$, and $\operatorname{dist}(x, -x) > \operatorname{diam}(E)$, a contradiction. Thus Equation 1 holds and the isodiametric inequality follows trivially for sets E symmetric about a point.

To prove it for general E, our next goal is to introduce a symmetric transformation that preserves $\mathcal{L}^n$ measure and does not increase the diameter: namely the Steiner symmetrization.

Definition 2.21 (Steiner Symmetrization). Let $E \subset \mathbb{R}^n$, and fix a unit vector $v \in \mathbb{S}^{n-1}$. Let P_v denote the hyperplane through the origin orthogonal to v. For each line L parallel to v, write point of intersection as $b_L = L \cap P_v$. The slice $E \cap L$ is replaced by the closed interval in L, centered at b_L, that has the same one-dimensional Lebesgue measure as $E \cap L$. The union of all such intervals, as L ranges over lines parallel to v, is called the

Steiner symmetrization of E with respect to v, denoted $S_v(E)$. More formally, we can write

$$S_v(E) := \bigcup_{\substack{L \| v \\ E\cap L \neq \emptyset}} \left\{ b_L + tv \;\middle|\; |t| \leq \tfrac{1}{2}\,\mathcal{H}^1(E\cap L) \right\},$$

where $b_L = P_v \cap L$, and $\mathcal{H}^1$ denotes one-dimensional Hausdorff measure. Note that $\mathcal{H}^1$ coincides with $\mathcal{L}^1$ in the subspace $L \simeq \mathbb{R}$ by Proposition 2.10.

Proposition 2.22. *Let $E \subset \mathbb{R}^n$ and let $v \in \mathbb{S}^{n-1}$ be a unit vector.*

(i) $\operatorname{diam}\big(S_v(E)\big) \leq \operatorname{diam}(E)$.

(ii) If E is $\mathcal{L}^n$ measurable then $S_v(E)$ is also $\mathcal{L}^n$ measurable and

$$\mathcal{L}^n\big(S_v(E)\big) = \mathcal{L}^n(E).$$

Proof. For $b \in P_v$ denote $L_b := b + \mathbb{R}v$, and define

$$\alpha(b) := \inf\{t \in \mathbb{R} :\; b + tv \in E\}, \qquad \beta(b) := \sup\{t \in \mathbb{R} :\; b + tv \in E\},$$

interpreting $\alpha(b) = +\infty,\; \beta(b) = -\infty$ when $E \cap L_b = \varnothing$. Note that for any b we always have

$$\mathcal{H}^1(E \cap L_b) \leq \beta(b) - \alpha(b),$$

because the line segment joining $\beta(b)$ and $\alpha(b)$ must have Lesbegue measure greater than $\mathcal{H}^1(E \cap L_b)$ by construction of S_v.

(i) If $\operatorname{diam} E = \infty$ the claim is trivial, so assume $\operatorname{diam} E < \infty$.
Let $x, y \in S_v(E)$. Then $x = b + t_x v,\; y = c + t_y v$ with $b, c \in P_v$ and $t_x, t_y \in \mathbb{R}$. By construction of $S_v(E)$ we have

$$|t_x| \leq \tfrac{1}{2}\,\mathcal{H}^1 E \cap L_b), \qquad |t_y| \leq \tfrac{1}{2}\,\mathcal{H}^1(E \cap L_c).$$

Without loss of generality assume (the other case is identical after swapping the roles of b and c)

$$\beta(c) - \alpha(b) \geq \beta(b) - \alpha(c) \tag{1}$$

and define

$$z_1 := b + \alpha(b)v, \qquad z_2 := c + \beta(c)v.$$

Note that as $\operatorname{diam}(E) < \infty$, all of $\alpha(b), \alpha(c), \beta(b), \beta(c)$ are finite. By definition of $\alpha(b), \beta(c)$, we have $z_1, z_2 \in \overline{E}$. Thus $|z_1 - z_2| \leq \operatorname{diam}(\overline{A}) = \operatorname{diam}(E)$. Our goal is to show $|x - y| \leq |z_1 - z_2|$, and as $x, y \in S_v(A)$ was arbitrary, this will prove $\operatorname{diam}\big(S_v(E)\big) \leq \operatorname{diam}(E)$.

By assumption 1,

$$\beta(c) - \alpha(b) \;\; \geq \;\; \tfrac{1}{2}\Big(\beta(b) - \alpha(b)\Big) + \tfrac{1}{2}\Big(\beta(c) - \alpha(c)\Big).$$

Since $\mathcal{H}^1(E \cap L_b) \leq \beta(b) - \alpha(b)$ and $\mathcal{H}^1(E \cap L_c) \leq \beta(c) - \alpha(c)$, it follows that

$$\beta(c) - \alpha(b) \;\; \geq \;\; \tfrac{1}{2}\mathcal{H}^1(E \cap L_b) + \tfrac{1}{2}\mathcal{H}^1(E \cap L_c).$$

Hence

$$\beta(c) - \alpha(b) \;\; \geq \;\; |t_x| + |t_y| \;\; \geq \;\; |t_x - t_y|.$$

Therefore

$$|x - y|^2 = |b - c|^2 + |t_x - t_y|^2 \;\; \leq \;\; |b - c|^2 + (\beta(c) - \alpha(b))^2 = |z_1 - z_2|^2.$$

(ii) As $\mathcal{L}^n$ is rotationally invariant, rotate coordinates so that $v = e_n$ and $P_v = \mathbb{R}^{n-1}$. For $b \in \mathbb{R}^{n-1}$ write $f(b) := \mathcal{H}^1(E \cap L_b)$. If E is $\mathcal{L}^n$ measurable, then by Fubini Tonelli theorem and $\mathcal{H}^1 = \mathcal{L}^1$ on lines L_b) (by Proposition 2.10, we have:

$$\begin{aligned}
\mathcal{L}^n(E) &= \int_{\mathbb{R}^{n-1}} \Big(\int_{\mathbb{R}} \chi_E(b, c)\, d\mathcal{L}^1(c) \Big)\, d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1(E \cap L_b)\, d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} f(b)\, db.
\end{aligned}$$

By Fubini, the marginal distribution f is $\mathcal{L}^{n-1}$ measurable. Define

$$g : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}, \qquad g(b, t) := \tfrac{1}{2} f(b) - |t|.$$

Since f is measurable and $(b, t) \mapsto |t|$ is continuous, g is measurable. Moreover

$$S_v(E) = \{(b, t) \in \mathbb{R}^{n-1} \times \mathbb{R} : \; |t| \leq \tfrac{1}{2} f(b)\} = \{(b, t) : g(b, t) \geq 0\}.$$

Thus $S_v(E)$ is a $\mathcal{L}^n$ measurable. Another application of Fubini on $\chi_{S_v(E)}$ yields:

$$L^n(S_v(E)) = \int_{\mathbb{R}^{n-1}} f(b)\, db,$$

thus proving (ii). $\square$

Lemma 2.23. *Let $v \in \mathbb{S}^{n-1}$ and let $R : \mathbb{R}^n \to \mathbb{R}^n$ be an isometry such that $R(v) = \pm v$. Then for every $\mathcal{L}^n$ measurable set $E \subset \mathbb{R}^n$, we have*

$$S_v\Big(R(E)\Big) = R\Big(S_v(E)\Big).$$

In particular, any reflection across a hyperplane that fixes the direction v (up to sign) commutes with S_v.

Proof. Since R is orthogonal and $R(v) = \pm v$, we have $R(P_v) = P_v$, and R maps each line $L_b := b + \mathbb{R}v$ (with $b \in P_v$) onto the line $L_{R(b)}$.
Fix $b \in P_v$. As $\mathcal{L}^1 = \mathcal{H}^1$ and $\mathcal{L}^1$ is invariant under isometries, we obtain

$$\mathcal{H}^1\Big(R(E) \cap L_{R(b)}\Big) = \mathcal{H}^1\Big(R(E \cap L_b)\Big) = \mathcal{H}^1\Big(E \cap L_b\Big).$$

By the construction of Steiner symmetrization, the slice of $S_v(R(E))$ along $L_{R(b)}$ is the centred segment in $L_{R(b)}$ of length $\mathcal{H}^1(R(E) \cap L_{R(b)})$, while the slice of $S_v(E)$ along L_b is the centred segment in L_b of length $\mathcal{H}^1(E \cap L_b)$. Applying R to the latter slice yields precisely the former slice because R sends the centre $b \in P_v$ to $R(b) \in P_v$ and preserves lengths on lines parallel to v.

Since this holds for every $b \in P_v$, the unions of the corresponding slices coincide: $S_v(R(E)) = R(S_v(E))$, as claimed. □

Proof of the Isodiametric Inequality 2.19. We are now in a position to prove the isodiametric inequality. If $\operatorname{diam}(E) = \infty$ the inequality is trivial, so assume $\operatorname{diam}(E) < \infty$.
Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. Define inductively

$$E^{(0)} := E, \qquad E^{(k)} := S_{e_k}\Big(E^{(k-1)}\Big) \quad (k = 1, \ldots, n),$$

Assume $E^{(k)}$ is symmetric about $P_{e_1}, ..., P_{e_k}$. By definition of $S_{e_{k+1}}$, $E^{(k+1)}$ is symmetric about $P_{e_{k+1}}$. Denote R_j to be the reflection across P_{e_j} for $j = 1$ to k. As $R_j(e_{k+1}) = e_{k+1}$, by Lemma 2.23, it follows that $R_j(E_{k+1}) = R_j(S_v(E_k)) = S_v(R_j(E_k)) = S_v(E_k) = E_{k+1}$, where $R_j(E_k) = E_k$ follows by our induction hypothesis that E_k is symmetric about P_{e_j}. Thus E_{k+1} is symmetric about $P_{e_1}, ..., P_{e_{k+1}}$

Thus, by induction, $E^{(n)}$ is symmetric about $P_{e_1}, ..., P_{e_n}$, i.e. $E^{(n)}$ is symmetric about the origin. By repeated application of Proposition 2.22, $\operatorname{diam}(E^{(n)}) \leq \operatorname{diam}(E)$ and $\mathcal{L}^n(E^{(n)}) = \mathcal{L}^n(E)$. Finally, by Remark 2.20, we discussed that the isodiametric inequality holds for sets symmetric about the origin. Thus, the isodiametric inequality holds for $E^{(n)}$, and thus E. □

Having established the isodiametric inequality, we can prove our main result:

Theorem 2.24 (Equivalence of $\mathcal{H}^n$ and $\mathcal{L}^n$)**.** *The n-dimensional Hausdorff outer measure coincides with the n-dimensional Lesbegue outer measure, i.e., $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$*

Proof. We prove the equivalence in two steps.

Step 1: $\mathcal{L}^n(E) \le \mathcal{H}^n(E)$. Let $E \subset \mathbb{R}^n$ and $\delta > 0$. Consider any countable cover $\{C_i\}$ of E with $\operatorname{diam}(C_i) < \delta$. By the isodiametric inequality,

$$\mathcal{L}^n(C_i) \le \alpha(n)\left(\frac{\operatorname{diam}(C_i)}{2}\right)^n.$$

Summing over i gives

$$\mathcal{L}^n(E) \le \sum_i \mathcal{L}^n(C_i) \le \sum_i \alpha(n)\left(\frac{\operatorname{diam}(C_i)}{2}\right)^n.$$

Taking the infimum over all such covers yields

$$\mathcal{L}^n(E) \le \mathcal{H}^n_\delta(E),$$

and letting $\delta \to 0$ gives

$$\mathcal{L}^n(E) \le \mathcal{H}^n(E).$$

Step 2: $\mathcal{H}^n(E) \le C\,\mathcal{L}^n(E)$. Fix $\varepsilon > 0$ and cover E by a countable collection of cubes $\{Q_i\}$ with $\operatorname{diam}(Q_i) < \delta$ such that

$$\sum_i \mathcal{L}^n(Q_i) \le \mathcal{L}^n(E) + \varepsilon.$$

Let P_i denote the circumscribed ball of Q_i. By a simple computation relating volume of P_i and Q_i, we get

$$\alpha(n)\left(\frac{\operatorname{diam}(P_i)}{2}\right)^n = \mathcal{L}^n(P_i) = C\,\mathcal{L}^n(Q_i)$$

for some constant C depending only on n. Then $\{P_i\}$ covers E and

$$\sum_i \alpha(n)\left(\frac{\operatorname{diam}(P_i)}{2}\right)^n \le C\sum_i \mathcal{L}^n(Q_i) \le C(\mathcal{L}^n(E) + \varepsilon),$$

so $\mathcal{H}^n(E) \le C(\mathcal{L}^n(E) + \varepsilon)$. Letting $\varepsilon \to 0$ gives $\mathcal{H}^n(E) \le C\mathcal{L}^n(E)$.
In particular, this gives $\mathcal{H}^n \ll \mathcal{L}^n$.

Step 3: Refinement using balls. Let $\delta > 0$ and $\{Q_i\}$ be as in Step 2. For each cube Q_i, by Corollary 1.12, we can find a countable collection of balls disjoint $\{B_{ij}\}_{j=1}^\infty$ such that $\operatorname{diam}(B_{ij}) < \delta$ and

$$\mathcal{L}^n\Big(\operatorname{int}(Q_i) \setminus \bigcup_j B_{ij}\Big) = 0 \implies \mathcal{L}^n\Big(Q_i \setminus \bigcup_j B_{ij}\Big) = 0. \tag{1}$$

Now, noting that (1) and $\mathcal{H}^n << \mathcal{L}^n$ implies $\sum_i \mathcal{H}^n_\delta\Big(Q_i \setminus \bigcup_j B_{ij}\Big) = 0$ and $\mathcal{H}^n_\delta(B_{ij}) \le$

$\alpha(n)\Big(\frac{\operatorname{diam}(B_{ij})}{2}\Big)^n$ by definition of $\mathcal{H}^n_\delta$, we get:

$$\begin{aligned}
\mathcal{H}^n_\delta(E) \leq \sum_i \mathcal{H}^n_\delta(Q_i) &\leq \sum_i \mathcal{H}^n_\delta\Big(\bigcup_j B_{ij}\Big) + \sum_i \mathcal{H}^n_\delta\Big(Q_i \setminus \bigcup_j B_{ij}\Big) \\
&= \sum_i \mathcal{H}^n_\delta\Big(\bigcup_j B_{ij}\Big) \leq \sum_i \sum_j \mathcal{H}^n_\delta(B_{ij}) \\
&\leq \sum_i \sum_j \alpha(n)\Big(\frac{\operatorname{diam}(B_{ij})}{2}\Big)^n \\
&= \sum_i \sum_j \mathcal{L}^n(B_{ij}) = \sum_i \mathcal{L}^n(Q_i) < \mathcal{L}^n(E) + \epsilon
\end{aligned}$$

Letting $\varepsilon \to 0$ and $\delta \to 0$ concludes the proof:

$$\mathcal{H}^n(E) = \mathcal{L}^n(E).$$

□

3 The Area Formula

3.1 Hausdorff Measures Under Linear Transformations

Having developed a good understanding of Hausdorff measures, it is now possible to explore how it changes under linear transformations. We begin with a simple lemma:

Lemma 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be an isometry (i.e. $\langle f(x), f(y)\rangle = \langle x, y\rangle$ for all x, y and note that f need not be surjective). Then for every $s \geq 0$ and every $A \subset \mathbb{R}^m$,*

$$\mathcal{H}^s\Big(f(A)\Big) = \mathcal{H}^s(A).$$

Proof. This follows trivially from the definition of $\mathcal{H}^s$. This is because given a cover $\{C_j\}$ of A, $\{f(C_j)\}$ is a cover of $f(A)$, and $\operatorname{diam}(f(C_j)) = \operatorname{diam}(C_j)$ as f is an isometry. $\square$

Definition 3.2 (Jacobian)**.** Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.
If $n \leq m$, define $\|L\|^2 := \det\Big(L^T L\Big)$, otherwise $\|L\|^2 := \det\Big(LL^T\Big)$.
If $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point $x \in U$, its *Jacobian at x* is defined by $Jf(x) := ||Df(x)\|$

Remark 3.3. Let $L = U\Sigma V^T$ be the singular value decomposition of L, where $U \in O(m)$, $V \in O(n)$, and Σ is the $m \times n$ diagonal matrix whose diagonal entries are the singular values $\sigma_1, \dots, \sigma_n$. Then:

$$||L|| = \det(L^t L)^{\frac{1}{2}} = \prod_{i=1}^{n} \sigma_i$$

As a corollary, $\sigma_n \leq ||L|| \leq \sigma_1$

Theorem 3.4. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map with $n \leq m$. For every set $A \subset \mathbb{R}^n$ one has $\mathcal{H}^n\Big(L(A)\Big) = ||L\|\ \mathcal{L}^n(A)$.*

Proof. Let $L = U\Sigma V^T$ be the singular value decomposition of L as in Remark 3.3. Writing $D = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$, we have $L = U \circ \iota \circ D \circ V^T$, where $\iota : \mathbb{R}^n \hookrightarrow \mathbb{R}^m$ is the canonical inclusion.

By Lemma 2.15, as U, V, ι are isometries,

$$\mathcal{H}^n(L(A)) = \mathcal{H}^n(D(A)) = \mathcal{L}^n(D(A))$$

and the last equality follows as $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$. Recall that:

$$\mathcal{L}^n(B) := \inf\Big\{\sum_i \operatorname{vol}(R_i) : \ B \subset \bigcup_i R_i,\ \ R_i \text{ rectangles}\Big\},$$

Thus, it is easy to see $\mathcal{L}^n(D(A)) = (\prod_{i=1}^n \sigma_i)\mathcal{L}^n(A)$ (Note that $\sigma_i \geq 0$). This is because for any rectangle R, the image $D(R)$ is still a rectangle, with its i-th side being scaled by σ_i; therefore $\mathcal{L}^n(D(R)) = (\prod_{i=1}^n \sigma_i)\mathcal{L}^n(R)$. Thus, given a cover $\{R_i\}$ of A, $\{D(R_i)\}$ is a cover of $D(A)$, and taking infimum over all such covers of rectangles of A, we get the inequality $\mathcal{L}^n(D(A)) \leq \det(D)\ \mathcal{L}^n(A)$. If $\det(D) = 0$, we are done, and if not, we get the reverse inequality with D^{-1}

Finally, since the numbers σ_i^2 are the eigenvalues of $L^T L$, we obtain

$$\mathcal{H}^n(L(A)) = \sqrt{\det(L^T L)}\, \mathcal{L}^n(A) = ||L\|\ \mathcal{L}^n(A). \qquad \square$$

Remark 3.5. The proof above which uses SVD avoids appealing directly to the formula $\mathcal{L}^n(T(A)) = |\det T|\, \mathcal{L}^n(A)$ for a linear map $T : \mathbb{R}^n \to \mathbb{R}^n$, which itself either relies on the classical change of variables theorem (whose derivation we aim to give as a corollary of the area formula) or on a lengthy measure-theoretic argument.

We conclude this section by observing a conveninet way of computing the Jacobian which shall be useful in the proof of the area formula.

Lemma 3.6 (Cauchy–Binet Formula). *Let A be an $n \times m$ matrix and B an $m \times n$ matrix over a field k. Then*

$$\det(AB) = \sum_{I \in S} \det\left(A_{\cdot,I}\right)\ \det\left(B_{I,\cdot}\right),$$

where:

- *$[m] = \{1, 2, ..., m\}$ and S the set of all n-combinations of $[m]$, with each $s \in S$ arranged in increasing order.*
- *$A_{\cdot,I}$ is the $n \times n$ submatrix of A formed by the columns indexed by I,*
- *$B_{I,\cdot}$ is the $n \times n$ submatrix of B formed by the rows indexed by I.*

Reference. A proof can be found in [6]. $\square$

Corollary 3.7 (Formula for Computing Jacobian).
Let $n \leq m$. If $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point $x \in U$, then

$$Jf(x)^2 = \det(Df(x)^T Df(x)) = \sum_{I \in S} \det\left(Df(x)_{\cdot,I}\right)^2$$

3.2 Properties of Lipschitz Maps

Next, our objective is to study some useful properties of Lipschitz. We start with a simple lemma:

Lemma 3.8. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and $0 \leq s < \infty$. Then for all $A \subset \mathbb{R}^n$,*

$$\mathcal{H}^s(f(A)) \leq \mathrm{Lip}(f)^s\, \mathcal{H}^s(A)$$

where

$$\mathrm{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} < \infty$$

Proof. This follows immediately from the definition of Hausdorff measure. Given a cover (U_i) of A, $(f(U_i))$ forms a cover of $f(A)$, and $\operatorname{diam}(f(U_i)) \leq Lip(f)\, diam(U_i)$, from which the desired inequality follows $\square$

Corollary 3.9. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and $A \subset \mathbb{R}^n$ such that $\mathcal{L}^n(A) = 0$. Then $\mathcal{H}^n(f(A)) = 0$*

Remark 3.10. The Lipschitz hypothesis is essential. A continuous map may send a Lebesgue null set to a set of positive $\mathcal{H}^n$-measure. For example, the *Cantor function* $c : [0,1] \to [0,1]$ is the unique continuous, nondecreasing map with $c(0) = 0$, $c(1) = 1$, and constant on each interval removed in the construction of the Cantor set. It maps the Cantor set $C \subset [0,1]$ (which satisfies $\mathcal{L}^1(C) = 0$) onto $[0,1]$, so $\mathcal{H}^1(c(C)) > 0$. Thus one cannot replace "Lipschitz" by "continuous" in Corollary 3.9.

Lemma 3.11 (Extending Lipschitz Functions)**.** *Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz map. Then there exists an extension $\tilde{f}$ to $\mathbb{R}^n$ of f such that $\tilde{f}(x) = f(x)$ $\forall x \in A$, and* $\operatorname{Lip}(\tilde{f}) \leq \sqrt{m}\operatorname{Lip}(f)$.

Proof. We first consider the case $m = 1$. Define, for $x \in \mathbb{R}^n$,

$$\tilde{f}(x) := \inf_{y \in A} \Big\{ f(y) + \operatorname{Lip}(f)\,|x - y| \Big\}.$$

Clearly, $\tilde{f}|_A = f$. For any $x_1, x_2 \in \mathbb{R}^n$, we have for all $y \in A$,

$$f(y) + \operatorname{Lip}(f)\,|x_1 - y| \leq f(y) + \operatorname{Lip}(f)\,|x_2 - y| + \operatorname{Lip}(f)\,|x_1 - x_2|,$$

taking the infimum over $y \in A$ on both sides gives

$$\tilde{f}(x_1) \leq \tilde{f}(x_2) + \operatorname{Lip}(f)\,|x_1 - x_2| \implies \tilde{f}(x_1) - \tilde{f}(x_2) \leq \operatorname{Lip}(f)\,|x_1 - x_2|.$$

Similarly,

$$\tilde{f}(x_2) - \tilde{f}(x_1) \leq \operatorname{Lip}(f)\,|x_1 - x_2|,$$

and thus $\operatorname{Lip}(\tilde{f}) = \operatorname{Lip}(f)$.

For a general m, write $f = (f_1, \ldots, f_m)$ and apply the previous construction to each component f_j. Then define

$$\tilde{f} := (\tilde{f}_1, \ldots, \tilde{f}_m),$$

which satisfies $\tilde{f}|_A = f$. Moreover,

$$|\tilde{f}(x_1) - \tilde{f}(x_2)|^2 = \sum_{j=1}^{m} |\tilde{f}_j(x_1) - \tilde{f}_j(x_2)|^2 \leq \sum_{j=1}^{m} \operatorname{Lip}(f)^2 |x_1 - x_2|^2 = m\,\operatorname{Lip}(f)^2 |x_1 - x_2|^2,$$

and thus

$$\operatorname{Lip}(\tilde{f}) \leq \sqrt{m}\,\operatorname{Lip}(f),$$

as desired. $\square$

In light of the lemma, we can assume, WLOG, our Lipschitz maps are defined on $\mathbb{R}^n$

Lemma 3.12. *Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz. Then f is differentiable a.e. w.r.t. $\mathcal{L}$.*

Proof. f Lipschitz $\Rightarrow$ f absolutely continuous $\Rightarrow$ f of bounded variation
$\Rightarrow$ $f = g - h,$ where g, h monotononically increasing functions
$\Rightarrow$ f differentiable a.e. as monotonic functions are differentiable a.e. $\square$

The next theorem, due to Rademacher, extends this result for Lipschitz functions defined on $\mathbb{R}^n$. Note that the above proof will not go through as absolute continuity is a notion only defined on $\mathbb{R}$

Theorem 3.13 (Rademacher's Theorem)**.** *Every Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable almost everywhere w.r.t. $\mathcal{L}^n$.*

Proof. It suffices to prove the theorem for $m = 1$, as the general case follows by applying the argument to each component separately.

Step 1: Set where Directional Derivative Exists has Measure 0 Fix $v \in S^{n-1}$. For $x \in \mathbb{R}^n$, define the upper and lower directional derivatives as

$$\overline{D}_v f(x) := \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t}, \qquad \underline{D}_v f(x) := \liminf_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Since f is Lipschitz, for all $x \in \mathbb{R}^n$ we have

$$-\operatorname{Lip}(f) \;\leq\; \underline{D}_v f(x) \;\leq\; \overline{D}_v f(x) \;\leq\; \operatorname{Lip}(f).$$

Thus the upper and lower directional derivatives alwaus exists, and the set of points where the directional derivative does not exist is

$$E := \{x \in \mathbb{R}^n : \underline{D}_v f(x) < \overline{D}_v f(x)\}.$$

Since the limsup and liminf can be written as limits over rational $t \neq 0$, and the difference quotient $\frac{f(x+tv)-f(x)}{t}$ is obviously Borel measurable as f Lipchitz, the functions $x \mapsto \overline{D}_v f(x)$ and $x \mapsto \underline{D}_v f(x)$ are Borel measurable. Hence, E is a Borel (and therefore $\mathcal{L}^n$) measurable set.

For each line L_x parallel v passing through x, the restriction $f|_{L_x}$ is a Lipschitz function of one variable. By Lemma 3.12 which proved the one-dimensional case, $f|_{L_x}$ is differentiable $\mathcal{H}^1$-almost everywhere along L_x. Hence for every line L_x, we have

$$\mathcal{H}^1(E \cap L_x) = 0.$$

Now consider a rotation $T \in O(n)$ sending $v \mapsto e_1$. As rotations preserve Lesbegue measure, $\mathcal{H}^1(T(E) \cap L'_x)$ where L'_x is line parallel to e_1 passing through x. Moreover, $T(E)$ is still Lebesgue measurable, and by Fubini's theorem,

$$\mathcal{L}^n(T(E)) = \int_{\mathbb{R}^n} \chi_{T(E)}(x)\,dx = \int_{\mathbb{R}^n} \left(\int_{T(E)\cap L'_{x_1}} dx_1 \right) dx_2 \cdots dx_n = \int_{\mathbb{R}^{n-1}} \mathcal{H}^1(T(E) \cap L'_{x_1})\,dx_2 \cdots dx_n = 0.$$

Thus $\mathcal{L}^n(E) = 0$, i.e. the set of points where the directional derivative along v fails to exist has measure zero.

Step 2: Directional derivatives in terms of partial derivatives. By step 1, $\nabla f = D_{e_1} f, \cdots, D_{e_n} f$ exists a.e. Fix $v \in S^{n-1}$, and let $\phi \in C_c^\infty(\mathbb{R}^n)$. Removing the measure 0 set where $D_v f$ and ∇f don't exist, consider the difference quotient along v:

$$\int_{\mathbb{R}^n} \frac{f(x+tv) - f(x)}{t}\,\phi(x)\,dx = -\int_{\mathbb{R}^n} f(x)\,\frac{\phi(x) - \phi(x-tv)}{t}\,dx.$$

Since f is Lipschitz,

$$\left| \frac{f(x+tv) - f(x)}{t} \right| \leq \mathrm{Lip}(f),$$

so the Dominated Convergence Theorem applies. Letting $t \to 0$, and using integration by parts, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\,\phi(x)\,dx &= -\int_{\mathbb{R}^n} f(x)\,D_v\phi(x)\,dx \\
&= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x)\,\partial_{x_i}\phi(x)\,dx. \\
&= \sum_{i=1}^n v_i \int_{\mathbb{R}^n} \partial_{x_i} f(x)\,\zeta(x)\,dx \\
&= \int_{\mathbb{R}^n} \big(v \cdot \nabla f(x)\big)\,\zeta(x)\,dx.
\end{aligned}$$

Since this equality holds for all test functions $\phi \in C_c^\infty(\mathbb{R}^n)$, we conclude in the sense of distributions. Moreover, as $D_v f,\ v \cdot \nabla f \in L^1_{loc}(\mathbb{R}^n)$, we conclude

$$D_v f(x) = v \cdot \nabla f(x) \quad \text{for a.e. } x \in \mathbb{R}^n.$$

Step 4: Differentiability at almost every point. Let $\{v_k\}_{k=1}^\infty$ be a countable dense subset of S^{n-1}. Define

$$E_k := \{x \in \mathbb{R}^n : D_{v_k} f(x) \text{ exists and } D_{v_k} f(x) = v_k \cdot \nabla f(x)\}, \quad E := \bigcap_{k=1}^\infty E_k.$$

As $\mathcal{L}^n(\mathbb{R}^n \setminus E_k) = 0\ \forall k$, we have $\mathcal{L}^n(\mathbb{R}^n \setminus E) = 0$. We'll show f is differentiable on $\mathbb{R}^n \setminus E$. For that, we first show $\nabla f(x) \cdot v$ is indeed the directional derivative $\forall v \in S^{n-1}$, not just those in the countabel dense subset. For $v \in S^{n-1}, x \in \mathbb{R}^n \setminus E$, and $t \neq 0$, let

$$A(x,v,t) = \frac{f(x+tv) - f(x)}{t} - \nabla f(x) \cdot v$$

Note that $|\nabla f(x)|^2 = \sum D_{e_i}(x)^2 \leq n \operatorname{Lip}(f)^2$. Thus,

$$\begin{aligned}
|A(x,v,t) &\leq |A(x,v_k,t)| + |A(x,v,t) - A(x,v_k,t)| \\
&\leq |A(x,v_k,t)| + |\frac{f(x+tv) - f(x+tv_k)}{t}| + |\nabla f(x) \cdot (v - v_k)| \\
&\leq |A(x,v_k,t)| + \operatorname{Lip}(f)|v - v_k| + \sqrt{n}\,|v - v_k| \\
&= |A(x,v_k,t)| + (1+\sqrt{n}) \operatorname{Lip}(f)|v - v_k|
\end{aligned}$$

Thus taking t small enough and v_k close enough, we make the first and second term resp. arbitrarily small. Thus $D_v f(x) = \nabla f(x) \cdot v \quad \forall v \in S^{n-1}$. To conclude the proof, note that:

$$\begin{aligned}
&\lim_{y \to x} \frac{|f(y) - f(x) - \nabla f(x) \cdot (y - x)|}{|y - x|} \\
&= \lim_{t \to 0} \frac{|f(x+tv) - f(x) - t\,\nabla f(x) \cdot (v)|}{t} \qquad (\text{ taking } v = \frac{y-x}{|y-x|} \in S^{n-1},\ t = |y-x|) \\
&= 0
\end{aligned}$$

This shows that f is differentiable almost everywhere, completing the proof. □

It follows that the Jacobian Jf is well defined a.e. for Lipschitz maps, which is sufficient for our purpose in making sense of the LHS of the area formula in 0.2

Definition 3.14 (Operator norm and co-norm)**.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.

- The *operator norm* is
$$\|T\|_{\mathrm{op}} := \sup_{|x|=1} |Tx|.$$
- The *co-norm* is
$$\|T\|_{\mathrm{co}} := \inf_{|x|=1} |Tx|.$$

Remark 3.15. Geometrically, $\|T\|_{\mathrm{op}}$ measures the maximal stretching of T, while $\|T\|_{\mathrm{co}}$ measures the minimal stretching. Let σ_1 and σ_r be the largest and smallest singular value resp. The following are easy to check properties:

- $\|T\|_{\mathrm{op}} = \mathrm{Lip(T)} = \sigma_1$
- $\|T\|_{\mathrm{co}} = \sigma_r$

- If T is bijective, $\|T\|_{\text{co}} = 1/\|T^{-1}\|_{\text{op}}$.
- If O is an isometry, $\|O \circ T\|_{\text{op}} = \|T\|_{\text{op}}$

Lemma 3.16. *Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuous and let*

$$D := \{x \in \mathbb{R}^n : \ g \textit{ is differentiable at } x\}.$$

Then the map $D \ni x \mapsto Dg(x) \in \mathbb{R}^{m\times n}$ is Borel.

Proof. Fix a sequence $(r_k) \subset \mathbb{Q}_{>0}$ with $r_k \downarrow 0$. For each k define the matrix of difference quotients

$$M_k(x)_{ij} := \frac{g_i(x + r_k e_j) - g_i(x)}{r_k}, \qquad x \in \mathbb{R}^n,$$

whose entries are continuous in x. Hence $M_k : \mathbb{R}^n \to \mathbb{R}^{m\times n}$ is Borel for each k.
On the differentiability set D, we have

$$Dg(x) = \lim_{k\to\infty} M_k(x),$$

so Dg is the pointwise limit of Borel functions on D, and therefore Borel.

□

We conclude this section with the most important ingredient in proving the area formula. The idea, as we shall see in the next section, is to find a cover E_i of $E = \{x : Jf(x) > 0\}$ (i.e. Df has full rank which implies local injectivity of f), such that f can be "approximated" by a linear injective map L_i in each E_i in a suitable way. The motivation is that we understand well how Hausdorff measure behaves under linear maps, while the global behavior of a Lipschitz map may vary wildly; locally, its behavior is controlled. This is made precise in the next theorem:

Theorem 3.17 (Linearization of Lipschitz Functions)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz map with $n \leq m$, and let*

$$E = \{x \in \mathbb{R}^n : f \textit{ is differentiable at } x \textit{ and } Jf(x) > 0\}.$$

Fix $t > 1$. Then there exists a countable cover of E by Borel sets $\{E_i\}$ such that:

1. *f is injective on each E_i;*
2. *for each i, there exists a symmetric automorphism $S_i : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$\mathrm{Lip}\Big(f|_{E_i} \circ S_i^{-1}\Big) < t, \qquad \mathrm{Lip}\Big(S_i \circ (f|_{E_i})^{-1}\Big) < t;$$

3. *and for all $x \in E_i$,*

$$t^{-n}|\det S_i| < Jf(x) < t^n|\det S_i|.$$

Proof. The argument follows a standard density approach. Let $\varepsilon > 0$ be such that

$$t^{-1} + \varepsilon < 1 < t - \varepsilon.$$

Let $\mathrm{Sym}(\mathbb{R}^n)$ denote the symmetric maps in $\mathbb{R}^n$, and $\mathrm{Sym}^*(\mathbb{R}^n) = \mathrm{Sym}(\mathbb{R}^n) \cap \mathrm{GL}(n, \mathbb{R})$.Since E and $\mathrm{Sym}^*(\mathbb{R}^n)$ are separable, there exist countable dense subsets $F \subset E$ and $A \subset \mathrm{Sym}^*(\mathbb{R}^n)$.
For $y \in F$, $n \in \mathbb{N}$, and $S \in A$, define

$$E(y, S, n) = \left\{ x \in B(y, 1/n) : \begin{array}{l} \text{(i) } \|Dg(x)\|_{\mathrm{op}} < t - \varepsilon, \quad \|Dg(x)\|_{\mathrm{co}} > t^{-1} + \varepsilon, \\ \text{(ii) } \dfrac{|g(z) - g(x) - Dg(x)(z - x)|}{|z - x|} < \varepsilon \quad \forall z \in B(x, 2/n) \end{array} \right\} \tag{1}$$

where $g = f \circ S^{-1}$, $\|\cdot\|_{\mathrm{op}}$ is the operator norm, and $\|\cdot\|_{\mathrm{co}}$ is the co-norm.
The idea is that condition (i) ensures $Df(x) \circ S^{-1}$ behaves almost like an isometry, meaning S captures how $Df(x)$ distorts volume, while condition (ii) guarantees that, in a neighborhood of x, the differential $Df(x)$ provides a good linear approximation of f.

It remains to show that the countable cover $\{E(y, S, n)\}_{y \in F,\, S \in A,\, n \in \mathbb{N}}$ satisfies the requirements of the theorem. Note that, by the polar decomposition, $Df(x) = O_x \circ U_x$, where $O_x : \mathbb{R}^n \to \mathbb{R}^m$ is an orthogonal map (linear isometry) and $U_x \in \mathrm{Sym}(\mathbb{R}^n)$ is symmetric $\forall x \in E$ Since $Jf(x) > 0$, the derivative $Df(x)$ has full rank, hence $U_x \in \mathrm{Sym}^*(\mathbb{R}^n)$

Claim. $E(y, S, n)$ is Borel.

Proof. This follows simply by Lemma 3.16 which says Dg is Borel on E. To see this, first note that by the continuity of the operator norm and Borel measurability of Dg, the compositions $x \mapsto \|Dg(x)\|_{\mathrm{op}}$ and $x \mapsto \|Dg(x)\|_{\mathrm{co}}$ are Borel on E. Therefore the set

$$\{x \in D : \ \|Dg(x)\|_{\mathrm{op}} < t - \varepsilon, \ \|Dg(x)\|_{\mathrm{co}} > t^{-1} + \varepsilon\} \text{ is Borel} \tag{2}$$

Next, define

$$\phi(x, z) := \frac{|g(z) - g(x) - Dg(x)(z - x)|}{|z - x|}, \quad z \neq x.$$

By Lemma 3.16, the map $(x, z) \mapsto \phi(x, z)$ is Borel on $E \times \Big(B(x, 2/n) \setminus \{x\}\Big)$. For each fixed $x \in D$, as g is continuous, the map $z \mapsto \phi(x, z)$ is continuous on $B(x, 2/n) \setminus \{x\}$, so for any countable dense $Q \subset B(x, 2/n)$,

$$\sup_{z \in B(x, 2/n) \setminus \{x\}} \phi(x, z) = \sup_{q \in Q} \phi(x, q).$$

The right-hand side is a countable supremum of Borel functions of x, hence $x \mapsto \sup_{z \in B(x, 2/n) \setminus \{x\}} \phi(x, z)$

is Borel on E. Consequently,

$$\{x \in E : \sup_{z \in B(x,2/n)\setminus\{x\}} \phi(x,z) < \varepsilon\} \text{ is Borel} \tag{3}$$

this set reflecting exactly condition (ii) of (1).
Finally, since $E(y,S,n)$ is the intersection of $B(y,1/n)$ with the Borel sets in (2) and (3) by definition, hence it is Borel. □

Claim. $\{E(y,S,n)\}$ *covers* E

Proof of Claim. To see this, first note that using properties in Remark 3.15,

$$\|Dg(x)\|_{\mathrm{op}} = ||Df(x)S^{-1}||_{\mathrm{op}} = ||O_x U_x S^{-1}||_{\mathrm{op}} = \|U_x S^{-1}\|_{\mathrm{op}}$$

and similarly,

$$\|Dg(x)\|_{\mathrm{co}} = \|U_x S^{-1}\|_{\mathrm{co}} = \frac{1}{\|SU_x^{-1}\|_{\mathrm{op}}}$$

Set

$$\varepsilon_1 := t - \varepsilon - 1, \qquad \varepsilon_2 := 1 - \frac{1}{t^{-1} + \varepsilon}.$$

By continuity of the inversion map on $\mathrm{Sym}(\mathbb{R}^n)$, there exists $\delta > 0$ such that

$$S \in B(U,\delta) \quad \Rightarrow \quad \|S^{-1} - U^{-1}\|_{\mathrm{op}} < \frac{\varepsilon_1}{\|U\|_{\mathrm{op}}}.$$

Choose S with

$$\|S - U\|_{\mathrm{op}} < \min\Big\{\delta,\ \varepsilon_2/\|U^{-1}\|_{\mathrm{op}}\Big\}.$$

We now estimate:

$$\begin{aligned}\|US^{-1}\|_{\mathrm{op}} &= \|\,U(S^{-1} - U^{-1}) + I\|_{\mathrm{op}} \\ &\le \|U\|_{\mathrm{op}}\|S^{-1} - U^{-1}\|_{\mathrm{op}} + \|I\|_{\mathrm{op}} \\ &< \|U\|_{\mathrm{op}}\frac{\varepsilon_1}{\|U\|_{\mathrm{op}}} + 1 = 1 + \varepsilon_1 = t - \varepsilon,\end{aligned}$$

and

$$\begin{aligned}\|SU^{-1}\|_{\mathrm{op}} &= \|\,(S - U)U^{-1} + I\|_{\mathrm{op}} \\ &\le \|S - U\|_{\mathrm{op}}\|U^{-1}\|_{\mathrm{op}} + 1 < \varepsilon_2 + 1 = \frac{1}{t^{-1} + \varepsilon}.\end{aligned}$$

Taking reciprocals gives

$$\|US^{-1}\|_{\mathrm{co}} = \frac{1}{\|SU^{-1}\|_{\mathrm{op}}} > t^{-1} + \varepsilon.$$

Let $g = f \circ S^{-1}$. Since $Dg(x) = Df(x) \circ S^{-1} = O\,US^{-1}$, we obtain

$$\|Dg(x)\|_{\mathrm{op}} = \|US^{-1}\|_{\mathrm{op}} < t - \varepsilon, \qquad \|Dg(x)\|_{\mathrm{co}} = \|US^{-1}\|_{\mathrm{co}} > t^{-1} + \varepsilon.$$

Hence S satisfies condition (i) in (1) Ensuring condition (ii) follows simply by definition of differentiability by choosing a small enough n. Finally, by density of F in E, $\exists y \in F$ such that $x \in B(y, 1/n)$. Thus $x \in E(y, S, n)$.

□

Claim. $\forall x \in E(y, S, n)$, $(t^{-1} + \epsilon)^{-n} |\det S| < Jf(x) < (t - \epsilon)^n |\det S|$.

Proof. Let $x \in E(y, S, n)$. Given that $Dg(x) = Df(x) \circ S^{-1}$,

$$\begin{aligned} Jf(x)^2 &= \det\Big(Df(x)^T Df(x)\Big) \\ &= \det\Big(S^T Dg(x)^T Dg(x) S\Big) \\ &= \det(S^T)\, \det\Big(Dg(x)^T Dg(x)\Big)\, \det(S) \\ &= Jg(x)^2\, \det(S)^2. \end{aligned} \tag{4}$$

As $Jg(x)$ is bounded above and below by σ_1 and σ_r, the largest and smallest singular of $Dg(x)$ resp., and $(\|Dg(x)\|_{\mathrm{op}} < t - \varepsilon,\ \|Dg(x)\|_{\mathrm{co}} > t^{-1} + \varepsilon$, we have

$$t^{-1} + \varepsilon < Jg(x) < t - \varepsilon$$

. Multiplying by $|\det(S)|$ and using Equation (4), we obtain the desired claim □

Claim. Let $G = E(y, S, n)$*. Then* $\operatorname{Lip}\Big(f|_G \circ S^{-1}\Big) < t$*,* $\operatorname{Lip}\Big(S \circ (f|_G)^{-1}\Big) < t$*;, and* $f|_G$ *is injective.*

Proof. If $x, z \in G$ and $z \in G$, then $x, z \in B(y, 1/n)$ so in particular $z \in B(x, 2/n)$. By condition (ii) and the operator–/co–norm bounds in (i) of the defintion of $E(y, S, n)$ in (1), we have

$$\begin{aligned} |g(z) - g(x)| &\le |g(z) - g(x) - Dg(x)(z - x)| + |Dg(x)(z - x)| \\ &< \varepsilon\,|z - x| + \|Dg(x)\|_{\mathrm{op}}\,|z - x| \\ &\le \varepsilon\,|z - x| + (t - \varepsilon)|z - x| \\ &= t\,|z - x|. \end{aligned}$$

Since this holds for every $x, z \in G$ we get $\operatorname{Lip}(g|_G) \le t$.

For the lower bound, again using (ii) and the co–norm bound in (i),

$$\begin{aligned}|g(z)-g(x)| &\geq |Dg(x)(z-x)| - |g(z)-g(x)-Dg(x)(z-x)| \\ &\geq \|Dg(x)\|_{\mathrm{co}}\,|z-x| - \varepsilon|z-x| \\ &> (t^{-1}+\varepsilon)|z-x| - \varepsilon|z-x| \\ &= t^{-1}|z-x|.\end{aligned}$$

Hence g is injective in G, and replacing z, x by $g^{-1}(z)$ and $g^{-1}(x)$, we get $|z-x| \geq t^{-1} g_{|G}^{-1}(z) - g(x)|$. Thus $\mathrm{Lip}((g|_G)^{-1}) < t$. As $g|_G = f|_G \circ S^{-1}$, this proves the claim □

□

3.3 Area Formula

We can now work towards a proof of the area formula. Before that, we prove a few prerequisite lemmas.

Lemma 3.18. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz. If $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable, then $f(A)$ is $\mathcal{H}^n$-measurable in $\mathbb{R}^m$.*

Proof. We treat the case $\mathcal{L}^n(A) < \infty$ first. By inner regularity of Lebesgue measure, there exists an increasing sequence of compact sets $K_i \subset A$ such that $\mathcal{L}^n\Big(A \setminus \bigcup_{i=1}^\infty K_i\Big) = 0$. By Lemma 3.8,

$$\mathcal{H}^n\Big(f(A \setminus \bigcup_i K_i)\Big) \leq \mathrm{Lip}(f)^n\, \mathcal{L}^n\Big(A \setminus \bigcup_i K_i\Big) = 0.$$

Write

$$f(A) = f\Big(\bigcup_i K_i\Big) \cup \Big(f(A) \setminus f\Big(\bigcup_i K_i\Big)\Big).$$

The second set is contained in $f\Big(A \setminus \bigcup_i K_i\Big)$, hence has $\mathcal{H}^n$-measure 0. As $\mathcal{H}^n$ is complete (measures induced by outer measures are complete) and each $f(K_i)$ is compact (thus $\mathcal{H}^n$-measurable), $\bigcup_i f(K_i)$ is measurable, and therefore $f(A)$ is $\mathcal{H}^n$-measurable.

If $\mathcal{L}^n(A) = \infty$ write $A = \bigcup_{R\geq 1}(A \cap B(0,R))$. Each $A \cap B(0,R)$ has finite Lebesgue measure, so by the previous paragraph $f(A \cap B(0,R))$ is $\mathcal{H}^n$-measurable; hence $f(A) = \bigcup_{R\geq 1} f(A \cap B(0,R))$ is measurable. □

Lemma 3.19 (Measurability of the Multiplicity Function)**.** *Let $n \leq m$ and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz map. Let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$ measurable. Then the map*

$$y \longmapsto \mathcal{H}^0\Big(f^{-1}(y) \cap A\Big)$$

called the multiplicity function is $\mathcal{H}^n$-measurable on $\mathbb{R}^m$.

Proof. For $k \in \mathbb{N}$, let $\mathcal{F}_k$ be the family of half-open cubes

$$\mathcal{F}_k := \left\{ \prod_{i=1}^{n} \left[\frac{a_i}{k}, \frac{a_i + 1}{k} \right) : a_i \in \mathbb{Z} \right\}.$$

For each k, define

$$g_k(y) := \sum_{Q \in \mathcal{F}_k} \mathbf{1}_{f(A \cap Q)}(y), \qquad y \in \mathbb{R}^m.$$

By Lemma 3.18, every set $f(A \cap Q)$ is $\mathcal{H}^n$-measurable; hence g_k is $\mathcal{H}^n$-measurable. Fix $y \in \mathbb{R}^m$ and let $N(y) := \mathcal{H}^0\Big(f^{-1}(y) \cap A\Big)$. By construction, $g_k(y)$ counts how many cubes $Q \in \mathcal{F}_k$ contain at least one element of $f^{-1}(y) \cap A$. Thus

$$g_k(y) \le g_{k+1}(y) \le N(y),$$

so $(g_k(y))_k$ is monotone and bounded above by $N(y)$.

If $N(y) < \infty$, say $f^{-1}(y) \cap A = \{x_1, \ldots, x_N\}$, Hausdorffness of $\mathbb{R}^n$ yields k_0 such that for $k \ge k_0$ each x_i lies in a distinct cube of $\mathcal{F}_k$. Hence $g_k(y) = N(y)$ for $k \ge k_0$. If $N(y) = \infty$, then for every $M \in \mathbb{N}$ one can choose M distinct preimages and find k_0 so that each lies in a different cube for all $k \ge k_0$, giving $g_k(y) \ge M$; hence $\lim_{k\to\infty} g_k(y) = \infty = N(y)$. Consequently

$$\lim_{k\to\infty} g_k(y) = \mathcal{H}^0\Big(f^{-1}(y) \cap A\Big) =: g(y).$$

Since each g_k is $\mathcal{H}^n$-measurable and $N = \lim_{k\to\infty} g_k$, the multiplicity function

$$y \longmapsto \mathcal{H}^0\Big(f^{-1}(y) \cap A\Big)$$

is $\mathcal{H}^n$-measurable. □

We now begin proving the area formula stated in Theorem 0.2. Let f, A be as in Theorem 0.2

Theorem 3.20. *Suppose $A \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$. Then Theorem 0.2 holds.*

Proof. Let $\mathcal{F}_k$ be as in Lemma 3.19. As $A \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$, by the Linearization of Lipschitz Functions theorem we obtain a countable family $\{E_i\}$ as in Theorem 3.17. Without loss of generality we may assume the sets E_i are pairwise disjoint (by replacing each E_i with $E_i \setminus \bigcup_{j<i} E_j$). Then $E_i \subset E(y, S, n)$ for some y, S, n as constructed in Theorem 3.17

For each k define the $\mathcal{H}^n$ measurable function

$$g_k(y) := \sum_{Q \in \mathcal{F}_k} \sum_{i \ge 1} \chi_{f(A \cap Q \cap E_i)}(y), \qquad y \in \mathbb{R}^m,$$

By the same argument as in Lemma 3.19 we have, for every y,

$$g_1(y) \leq g_2(y) \leq \cdots \quad \text{and} \quad \lim_{k\to\infty} g_k(y) = \mathcal{H}^0\Big(f^{-1}(y) \cap A\Big) =: N(y).$$

Since each g_k is $\mathcal{H}^n$-measurable and $g_k \uparrow N$, the Monotone Convergence Theorem yields.

$$\int_{\mathbb{R}^m} g_k \, d\mathcal{H}^n \longrightarrow \int_{\mathbb{R}^m} N \, d\mathcal{H}^n \qquad (k \to \infty).$$

Now let $B := A \cap E_i \cap Q \subset E(y, S, n)$ for some y, S, n. Using the properties in Lemma 3.19, we obtain:

$$\begin{aligned}
t^{-2n}\,\mathcal{H}^n(f(B)) &\leq t^{-2n} H^n((fS^{-1})(S(B))) \\
&\leq t^{-n}\mathcal{H}^n(S(B)) \quad (\text{as } \operatorname{Lip}(fS^{-1}) \leq t) \\
&= t^{-n}\,|detS|\,H^n(B) \\
&= \int_B Jf(x)\,dx \\
&= t^n\,|detS|\,\mathcal{H}^n(B) \\
&= t^n\mathcal{H}^n(S(B)) \\
&= t^n H^n((Sf^{-1})(f(B))) \\
&= t^{2n} H^n(f(B)) \quad (\text{as } \operatorname{Lip}(Sf^{-1}) \leq t)
\end{aligned}$$

Thus, by MCT, summing over all pairs (Q, i) that define g_k gives

$$t^{-2n} \int_{\mathbb{R}^m} g_k(y)\,d\mathcal{H}^n(y) \leq \int_A Jf(x)\,dx \leq t^{2n} \int_{\mathbb{R}^m} g_k(y)\,d\mathcal{H}^n(y).$$

Letting $k \to \infty$ and using the MCT yet again yields

$$t^{-2n} \int_{\mathbb{R}^m} N(y)\,d\mathcal{H}^n(y) \leq \int_A Jf(x)\,dx \leq t^{2n} \int_{\mathbb{R}^m} N(y)\,d\mathcal{H}^n(y).$$

Finally let $t \downarrow 1$ to obtain the desired equality

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

$\square$

Lemma 3.21. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, then*

$$\|Df(x)\|_{\mathrm{op}} \leq \operatorname{Lip}(f) \qquad \textit{for a.e. } x \in \mathbb{R}^n \textit{ where } Df(x) \textit{ exists.}$$

Proof. This is standard: for any unit vector v,

$$\limsup_{h\to 0} \frac{|f(x+hv)-f(x)|}{|h|} \le \mathrm{Lip}(f),$$

so the operator norm of $Df(x)$, whenever it exists, is bounded by $\mathrm{Lip}(f)$. $\square$

Theorem 3.22. *Suppose $A \subset \{x \in \mathbb{R}^n : Jf(x) = 0\}$. Then Theorem 0.2 holds.*

Remark 3.23. Note that the proof boils down to showing $\mathcal{H}^n(f(A)) = 0$. A special case of the above Sard's Theorem when $n \le m$, which says that for f a C^1 function, $\mathcal{L}^m(f(B)) = 0$ where B is the set of critical point of f, i.e, $B = \{x : Df(x) \text{ does not have full rank}\} = \{x : Jf(x) = 0\}$. The case $n \ge m$ of Sard's Theorem will be a special of the Co-Area Formula, which we shall see in the end-semester report.

Proof of Theorem 3.22. Let $0 < \varepsilon < 1$ and define the auxiliary map

$$g(x) := \big(f(x),\, \varepsilon x\big), \qquad x \in \mathbb{R}^n.$$

Geometrically, g embeds the graph of f into $\mathbb{R}^{m+n}$, adding a small "vertical" component so that the Jacobian of g is strictly positive as shown below, allowing us to use Theorem 3.20.

By Corollary 3.7 (Cauchy–Binet formula for the Jacobian), the n-Jacobian of g is

$$Jg(x)^2 = \sum_I \big(\det M_I(x)\big)^2,$$

where $M_I(x)$ runs over the $n \times n$ minors of the matrix $Dg(x) = \big(Df(x) \mid \varepsilon I_n\big)$. Two easy bounds follow:

1. *Lower bound* One of the minors is the determinant of the block εI_n; hence

$$Jg(x) \;\ge\; \varepsilon^{2n} > 0.$$

2. *Upper bound* Every column of $Df(x)$ has norm at most $\mathrm{Lip}(f)$ by Lemma 3.21, while columns coming from εI_n have norm ε. Each term in the Cauchy–Binet sum either coincides with a minor of $Df(x)$ (giving $Jf(x)$) or contains at least one column with factor ε. Consequently,

$$Jg(x) \le Jf(x) + C\,\varepsilon^2 = C\,\varepsilon^2,$$

where the constant C depends only on n, m and $\mathrm{Lip}(f)$, not on x. (The power ε^2 reflects that any "mixed" minor uses at least two columns from εI_n.)

Assume first that $\mathcal{L}^n(A) < \infty$. Let $p : \mathbb{R}^{m+n} \to \mathbb{R}^m$ be the projection $p(y,z) = y$. Since $\mathrm{Lip}(p) = 1$,

$$\mathcal{H}^n\big(f(A)\big) = \mathcal{H}^n\big(p \circ g(A)\big) \le \mathcal{H}^n\big(g(A)\big) = \int_A Jg(x)\,dx \le C\,\varepsilon^2,$$

by Theorem 3.20 (which applies as $Jg(x) > 0$ and the bound on $Jg(x)$. Letting $\varepsilon \downarrow 0$ gives $\mathcal{H}^n(f(A)) = 0$.
If $\mathcal{L}^n(A) = \infty$, write $A = \bigcup_k (A \cap B_k)$ with B_k balls of finite measure and apply the above to each piece. □

It is now easy to see that the area formula in Theorem 0.2 follows easily:

Proof of Area Formula. Decompose $A = A_1 \cup A_2 \cup A_3$, where A_1 is the set where f is not differentiable ($\mathcal{L}^n(A_1) = 0$), $A_2 = \{x \in A : Jf(x) > 0\}$, and $A_3 = \{x \in A : Jf(x) = 0\}$. Applying Theorem 3.20 on A_2 and Theorem 3.22 on A_3, and observing that A_1 has measure 0 by Rademacher's Theorem 3.13, yields the area formula. □

Proof of the Generalized Change of Variables. The generalized change of variables formula stated in the introduction 0.3 follows easily from the area formula. To see this, observe that the equality holds trivially for a simple function g from the area formula. Next, for any g integrable, write $g = g^+ - g^-$, and apply MCT to get the desired result. □

We conclude this section with a powerful application from [7]

Theorem 3.24. *Let $M \subset \mathbb{R}^n$ be a C^1 (embedded) oriented k-dimensional submanifold, endowed with the Riemannian metric g induced by the Euclidean inner product of $\mathbb{R}^n$. Let $\mathcal{H}^k|_M$ denote the restriction of k-dimensional Hausdorff measure to M, and let $\mathrm{d}V_g$ be the Riemannian volume form on M (the unique k-form whose value on an oriented orthonormal basis is 1). Then $\mathcal{H}^k\big|_M$ agrees with volume form $\mathrm{d}V_g$, i.e.*

$$\mathcal{H}^k(E) = \mathrm{vol}(E) = \int_E \mathrm{d}V_g \qquad \textit{for every } \mathcal{H}^k \textit{ measurable subset } E \subset M$$

.

Proof. Choose a countable atlas $\{(U_i, \varphi_i)\}$ of oriented charts for M such that U_i is precompact and $\varphi_i : U_i \subset M \to V_i \subset \mathbb{R}^k$ a C^1 parametrization. As φ_i^{-1} is C^1 and V_i is precompact (as U_i is) , φ_i^{-1} is Lipschitz.
Fix a chart $\varphi = \varphi_i$, $U = U_i$ and assume $E \subset \varphi(U)$. By the area formula,

$$\mathcal{H}^k(E) = \mathcal{H}^k\Big(\varphi^{-1}(\varphi(E))\Big) = \int_{\varphi(E)} \sqrt{\det\Big(D\varphi^{-1}(x)^T D\varphi^{-1}(x)\Big)}\, dx. \tag{1}$$

Now, denoting $\varphi = (x_1, ..., x_k)$ and recalling the construction of the volume form $\mathrm{d}V_g$, we get,

$$\mathrm{d}V_g = \sqrt{\det(g_{ij})}\, dx_1 \wedge \cdots \wedge dx_k,$$
$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x_i}\bigg|_p, \frac{\partial}{\partial x_j}\bigg|_p \right\rangle.$$

Thus,

$$\mathrm{vol}(E) = \int_{\varphi(E)} \sqrt{\det(g_{ij})} \circ \varphi^{-1}(x)\, dx \tag{2}$$

Observing that $D\varphi^{-1}(x)(\frac{\partial}{\partial r_i}\Big|_x) = \frac{\partial}{\partial x_i}\Big|_{\varphi^{-1}(x)}$ where $(r_i)_{i=1}^k$ denotes the standard coordinate in $\mathbb{R}^k$, the integrands in (1) and (2) are equal, hence $\mathcal{H}^k(E) = \mathrm{vol}(E)$

Having established the equality $\mathcal{H}^k(\varphi(E)) = \int_{\varphi(E)} dV_g$ for every $\mathcal{H}^k$ measurable $E \subset U_i$ in a single chart, the global equality on M follows immediately. Indeed $\{\varphi_i(U_i)\}_i$ is a countable cover of M; writing

$$E = \bigcup_i \Big(E \cap \varphi_i(U_i)\Big)$$

and replacing the family by the disjoint pieces $\widetilde{E}_1 := E \cap \varphi_1(U_1)$, $\widetilde{E}_2 := E \cap \varphi_2(U_2) \setminus \varphi_1(U_1)$, etc., we obtain a partition of E into countably many $\mathcal{H}^k$ measurable sets each contained in a single chart image.

Then

$$\mathrm{vol}(E) = \sum \mathrm{vol}(\widetilde{E}_i) = \sum \mathcal{H}^k(\widetilde{E}_i) = \mathcal{H}^k(E)$$

completing the proof. $\square$

References

[1] Giovanni Alberti. A brief introduction to geometric measure theory. https://www.math.stonybrook.edu/~bishop/classes/math638.F20/Alberti_GMT_brief_intro.pdf, 2020. Lecture notes for Math 638, Stony Brook University.

[2] Stanford Mathematics Department. Borel regular & radon measures, 2013. Accessed: 2025-09-23.

[3] Lawrence C. Evans and Ronald F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, revised edition edition, 1992.

[4] Herbert Federer. *Geometric Measure Theory*. Berlin, Heidelberg, New York, Springer, 1969.

[5] A. A. Kirillov. A tale of two fractals. https://www2.math.upenn.edu/~kirillov/MATH480-F07/tf.pdf, 2007.

[6] Andrew Putman. The cauchy-binet formula. https://academicweb.nd.edu/~andyp/notes/CauchyBinet.pdf, 2011.

[7] Leon Simon. *Introduction to Geometric Measure Theory*. Stanford University, 2014. Available online: https://web.stanford.edu/class/math285/ts-gmt.pdf.