NOTES ON L. C. EVANS AND R. F. GARIEPY: *MEASURE THEORY AND FINE PROPERTIES OF FUNCTIONS*

A. D. WENDLAND

Notes on chapters 2, 3, and 5 of *Measure Theory and Fine Properties of Functions* by L. C. Evans and R. F. Gariepy. All references are from eg:measure [1] unless indicated otherwise.

Contents

---

1. General Measure Theory

1.1. Weak Convergence and Compactness for Radon Measures.

t1.9-1 **Theorem 1.1.1.** *Let $\mu, \{\mu_k\}_{k=1}^{+\infty}$ be Radon measures on $\mathbb{R}^n$. The following three statements are equivalent:*

(i) $\lim_{k\to+\infty} \int_{\mathbb{R}^n} f \, d\mu_k = \int_{\mathbb{R}^n} f \, d\mu$ *for all* $f \in \mathcal{C}_c(\mathbb{R}^n)$;
(ii) $\limsup_{k\to+\infty} \mu_k(K) \le \mu(K)$ *for each compact set* $K \subseteq \mathbb{R}^n$ *and* $\mu(U) \le \liminf_{k\to+\infty} \mu_k(U)$ *for each open set* $U \subseteq \mathbb{R}^n$;
(iii) $\lim_{k\to+\infty} \mu_k(B) = \mu(B)$ *for each bounded Borel set* $B \subseteq \mathbb{R}^n$ *with* $\mu(\partial B) = 0$.

Remark. *Recall that Radon measures on $\mathbb{R}^n$ are characterized by inner and outer regularity. Let $B \subseteq \mathbb{R}^n$ be a Borel set, and let $K \subseteq B \subseteq U$ with K compact and U open. If $\{\mu_k\}_{k=1}^{+\infty}$ is converging to μ in any sense, we should expect $\mu_k(K) \le \mu(K)$ for all $k \in \mathbb{N}$ and $\mu_k(U) \ge \mu(U)$ for all $k \in \mathbb{N}$. Conditions (ii) and (iii) tell us that this in fact holds up to a subsequence.*

Date: March 30, 2023.

Definition 1.1.1 (Weak Convergence of Radon Measures). *Let $\mu, \{\mu_k\}_{k=1}^{+\infty}$ be Radon measures on $\mathbb{R}^n$. We say that $\{\mu_k\}_{k=1}^{+\infty}$ converges weakly to μ, and write*

$$\mu_k \rightharpoonup \mu,$$

if

$$\lim_{k\to+\infty} \int_{\mathbb{R}^n} f \, d\mu_k = \int_{\mathbb{R}^n} f \, d\mu$$

for every $f \in \mathcal{C}_c(\mathbb{R}^n)$.

Proof. Assume first that (i) holds. Let $U \subseteq \mathbb{R}^n$ be open, and choose a compact set $K \subseteq U$. Next apply Urysohn's Lemma to choose a function $f \in \mathcal{C}_c(\mathbb{R}^n)$ such that

$$0 \le f \le 1, \quad \operatorname{supp}(f) \subseteq U, \quad \text{and} \quad f \equiv 1 \text{ on } K.$$

Then

$$\begin{aligned}\mu(K) = \int_K d\mu = \int_K f \, d\mu \le \int_{\mathbb{R}^n} f \, d\mu = \lim_{k\to+\infty} \int_{\mathbb{R}^n} f \, d\mu_k \le \liminf_{k\to+\infty} \int_U \, d\mu_k \\ = \liminf_{k\to\infty} \mu_k(U).\end{aligned}$$

Thus

$$\begin{aligned}\mu(U) &= \sup\{\mu(K) : K \text{ compact}, K \subseteq U\} \\ &\le \liminf_{k\to+\infty} \mu_k(U).\end{aligned}$$

This proves the second part of (ii). The first part is similar.

Next suppose that (ii) holds. Let $B \subseteq \mathbb{R}^n$ be a bounded Borel set, $\mu(\partial B) = 0$. Then by (ii),

$$\begin{aligned}\mu(B) = \mu(B^\circ) &\le \liminf_{k\to+\infty} \mu_k(B^\circ) \\ &\le \limsup_{k\to+\infty} \mu_k(\overline{B}) \\ &\le \mu(\overline{B}) \\ &= \mu(B).\end{aligned}$$

Since $\mu_k(B^\circ) = \mu_k(B) = \mu_k(\overline{B})$ for all $k \in \mathbb{N}$ since $\mu(\partial B) = 0$, it follows

$$\liminf_{k\to+\infty} \mu_k(B) = \limsup_{k\to+\infty} \mu_k(B).$$

Thus $\lim_{k\to+\infty} \mu_k(B)$ exists, and

$$\lim_{k\to+\infty} \mu_k(B) = \mu(B),$$

as required.

Finally assume that (iii) holds. Fix $\epsilon > 0$ and $f \in \mathcal{C}_c^+(\mathbb{R}^n)$. Let $R > 0$ be such that $\operatorname{supp}(f) \subseteq B(0, R)$ and $\mu(\partial B(0, R)) = 0$. Choose a partition

$$0 := t_0 < t_1 < \cdots < t_N = 2\|f\|_{L^\infty(\mathbb{R}^n)}$$

of $[0, 2\|f\|_{L^\infty(\mathbb{R}^n)}]$ such that $0 < t_i - t_{i-1} < \epsilon$, and $\mu(f^{-1}\{t_i\}) = 0$ for each $i = 1, \ldots, N$. Put $B_i := f^{-1}((t_{i-1}, t_i])$, $i = 2, \ldots, N$. Then $\mu(\partial B_i) = 0$ for each $i \geq 2$. Now

$$\begin{aligned}\sum_{i=2}^N t_{i-1}\mu_k(B_i) = \sum_{i=2}^N t_{i-1} \int_{B_i} d\mu_k &\leq \sum_{i=2}^N \int_{B_i} f\, d\mu_k \\ &\leq \int_{\mathbb{R}^n} f\, d\mu_k \\ &\leq \sum_{i=2}^N t_i\mu_k(B_i) + t_1\mu_k(B(0,R)),\end{aligned}$$

and

$$\begin{aligned}\sum_{i=2}^N t_{i-1}\mu(B_i) = \sum_{i=2}^N t_{i-1} \int_{B_i} d\mu &\leq \sum_{i=2}^N \int_{B_i} f\, d\mu \\ &\leq \int_{\mathbb{R}^n} f\, d\mu \\ &\leq \sum_{i=2}^N t_i\mu(B_i) + t_1\mu(B(0,R)).\end{aligned}$$

Thus (iii) implies

$$\begin{aligned}&\limsup_{k\to+\infty} \left| \int_{\mathbb{R}^n} f\, d\mu_k - \int_{\mathbb{R}^n} f\, d\mu \right| \\ &\leq \limsup_{k\to+\infty} \left| \left\{ \sum_{i=2}^N t_i\mu_k(B_i) + t_1\mu_k(B(0,R)) \right\} - \sum_{i=2}^N t_{i-1}\mu(B_i) \right| \\ &\leq \limsup_{k\to+\infty} \sum_{i=2}^N |t_i\mu_k(B_i) - t_{i-1}\mu(B_i)| + \limsup_{k\to+\infty} t_1\mu_k(B(0,R)) \\ &= \sum_{i=2}^N |t_i - t_{i-1}|\mu(B_i) + t_1\mu(B(0,R)) \\ &\leq 2\epsilon\mu(B(0,R)).\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, taking the limit at $\epsilon \to 0$ shows that

$$\limsup_{k\to+\infty} \left| \int_{\mathbb{R}^n} f\, d\mu_k - \int_{\mathbb{R}^n} f\, d\mu \right| = 0,$$

and hence

$$\lim_{k\to+\infty} \int_{\mathbb{R}^n} f\, d\mu_k = \int_{\mathbb{R}^n} f\, d\mu.$$

The proof is complete. □

t1.9-2 **Theorem 1.1.2** (Weak Compactness for Measures). *Let $\{\mu_k\}_{k=1}^{+\infty}$ be a sequence of Radon measures on $\mathbb{R}^n$ satisfying*

$$\sup_{k\in\mathbb{N}} \mu_k(K) < +\infty$$

for each compact set $K \subseteq \mathbb{R}^n$. Then there exists a subsequence $\{\mu_{k_j}\}_{j=1}^{+\infty}$ and a Radon measure μ on $\mathbb{R}^n$ such that

$$\mu_{k_j} \rightharpoonup \mu \quad \text{as } j \to +\infty.$$

Proof.

(i). Assume first that

$$\sup_{k \in \mathbb{N}} \mu_k(\mathbb{R}^n) < +\infty. \tag{1.1.1}$$

(ii). Let $\{f_k\}_{k=1}^{+\infty}$ be a countable dense subset of $\mathcal{C}_c(\mathbb{R}^n)$. Note that (1.1.1) implies that the sequence $\{\int_{\mathbb{R}^n} f_1 \, d\mu_j\}_{j=1}^{+\infty}$ is bounded, for

$$\left| \int_{\mathbb{R}^n} f_1 \, d\mu_j \right| \leq \int_{\mathbb{R}^n} |f_1| \, d\mu_j \leq \max_{x \in \operatorname{supp}(f)} |f(x)| \mu_j(\mathbb{R}^n) < +\infty.$$

Thus we may find a subsequence $\{\mu_j^1\}_{j=1}^{+\infty}$ and $a_1 \in \mathbb{R}$ such that

$$\int_{\mathbb{R}^n} f_1 \, d\mu_j^1 \to a_1 \quad \text{as} \quad j \to +\infty.$$

Continuing, we find subsequences $\{\mu_j^k\}_{j=1}^{+\infty}$ of $\{\mu_j^{k-1}\}_{j=1}^{+\infty}$ and numbers $a_k \in \mathbb{R}$ such that

$$\int_{\mathbb{R}^n} f_k \, d\mu_j^k \to a_k \quad \text{as} \quad j \to +\infty$$

for each $k \in \mathbb{N}$. Set $\nu_j := \mu_j^j$. Then

$$\int_{\mathbb{R}^n} f_k \, d\nu_j \to a_k \quad \text{as} \quad j \to +\infty$$

for all $k \in \mathbb{N}$, for if $j \geq k$, then $\nu_j = \mu_j^j \in \{\mu_j^k\}_{j=1}^{+\infty}$. Define $L(f_k) := a_k$, and note that L is linear and

$$|L(f_k)| \leq M \|f_k\|_{L^\infty(\mathbb{R}^n)}$$

by (1.1.1), where

$$M := \sup_{k \in \mathbb{N}} \mu_k(\mathbb{R}^n).$$

By the Hahn–Banach Theorem, L may be uniquely extended to a bounded linear functional $\overline{L}$ defined on all of $\mathcal{C}_c(\mathbb{R}^n)$. Then, by the Riesz Representation Theorem, there exists a unique Radon measure μ on $\mathbb{R}^n$ such that

$$\overline{L}(f) = \int_{\mathbb{R}^n} f \, d\mu$$

for all $f \in \mathcal{C}_c(\mathbb{R}^n)$.

(iii). Choose any $f \in \mathcal{C}_c(\mathbb{R}^n)$. Since $\{f_k\}_{k=1}^{+\infty}$ is dense in $\mathcal{C}_c(\mathbb{R}^n)$, there exists a subsequence $\{f_{k_i}\}_{i=1}^{+\infty}$ such that $f_i \to f$ uniformly. Fix $\epsilon > 0$ and then choose $i \in \mathbb{N}$ so large that

$$\|f_{k_i} - f\|_{L^\infty(\mathbb{R}^n)} < \frac{\epsilon}{4M}. \tag{1.1.2}$$

Next choose $J \in \mathbb{N}$ so that for all $j > J$,

$$\left| \int_{\mathbb{R}^n} f_{k_i} \, d\nu_j - \int_{\mathbb{R}^n} f_{k_i} \, d\mu \right| < \frac{\epsilon}{2}.$$

Then for any $j > J$, we have by (eq:1.9-2 1.1.2) and the Principle of Uniform Boundedness

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f\, d\nu_j - \int_{\mathbb{R}^n} f\, d\mu\right| &\leq \left|\int_{\mathbb{R}^n} f - f_{k_i}\, d\nu_j\right| + \left|\int_{\mathbb{R}^n} f_{k_i}\, d\nu_j - \int_{\mathbb{R}^n} f_{k_i}\, d\mu\right| + \\ &\qquad \left|\int_{\mathbb{R}^n} f_{k_i} - f\, d\mu\right| \\ &\leq \frac{\epsilon}{2} + \|f - f_{k_i}\|_{L^\infty(\mathbb{R}^n)}\nu_j(\mathbb{R}^n) + \|f - f_{k_i}\|_{L^\infty(\mathbb{R}^n)}\mu(\mathbb{R}^n) \\ &< \epsilon,\end{aligned}$$

as required.

(iv). In the general case that (eq:1.9-1 1.1.1) fails to hold, but

$$\sup_{k\in\mathbb{N}} \mu_k(K) < +\infty$$

for each compact set $K \subseteq \mathbb{R}^n$, we apply the above argument to the measures

$$\mu_k^l := \mu_k \llcorner \overline{B(0,l)}, \quad k, l = 1, 2, \dots,$$

and use a diagonalization argument. The proof is complete. □

For the remainder of this section, we assume that

(i) $U \subseteq \mathbb{R}^n$ is open;
(ii) $1 \leq p < +\infty$.

Definition 1.1.2 (Weak Convergence in $L^p(U)$)**.** *A sequence $\{f_k\}_{k=1}^{+\infty} \subset L^p(U)$ is said to converge weakly to $f \in L^p(U)$, written*

$$f_k \rightharpoonup f \quad \text{in } L^p(U),$$

if

$$\lim_{k\to+\infty} \int_U f_k g\, d\mathcal{L}^n = \int_U fg\, d\mathcal{L}^n$$

for each $g \in L^q(U)$, where p and q are conjugate exponents, $\frac{1}{p} + \frac{1}{q} = 1$, $1 < q \leq +\infty$.

t1.9-3 **Theorem 1.1.3** (Weak Compactness in L^p)**.** *Suppose that $1 < p < +\infty$. Let $\{f_k\}_{k=1}^{+\infty} \subseteq L^p(U)$ satisfying*

$$\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(U)} < +\infty.$$

Then there exists a subsequence $\{f_{k_j}\}_{j=1}^{+\infty}$ of $\{f_k\}_{k=1}^{+\infty}$ and a function $f \in L^p(U)$ such that

$$f_{k_j} \rightharpoonup f \quad \text{in } L^p(U) \quad \text{as } j \to +\infty.$$

Remark. *This assertion is in general false for $p = 1$. The key property here is reflexivity. Recall that $L^p(U)$ is reflexive if and only if $1 < p < +\infty$.*

Definition 1.1.3. *We denote by*

$$\nu := \mu \llcorner f$$

the signed measure with density f with respect to μ, that is, the signed measure

$$\nu(K) = \int_K f\, d\mu,$$

provided that this holds for all compact sets $K \subseteq \mathbb{R}^n$.

Proof.

(i). If $U \neq \mathbb{R}^n$, we extend each function f_k to $\mathbb{R}^n$ by setting $f_k = 0$ on $\mathbb{R}^n \setminus U$. This done, we may assume that $U = \mathbb{R}^n$. We may also assume that

$$f_k \geq 0 \quad \mathcal{L}^n - \text{a.e.},$$

for otherwise we could apply the following analysis to f_k^+ and f_k^-.

(ii). Define the Radon measures

$$\mu_k := \mathcal{L}^n \llcorner f_k, \quad k \in \mathbb{N}.$$

Then for each compact set $K \subseteq \mathbb{R}^n$, by Hölder's inequality, we have

$$\mu_k(K) = \int_K f_k \, d\mathcal{L}^n \leq \|f_k\|_{L^p(K)} \cdot \mathcal{L}^n(K)^{\frac{p-1}{p}} < +\infty,$$

and thus

$$\sup_{k \in \mathbb{N}} \mu_k(K) < +\infty.$$

Therefore, we may apply Theorem (1.1.2) to obtain a Radon measure μ on $\mathbb{R}^n$ and a subsequence

$$\mu_{k_j} \rightharpoonup \mu.$$

(iii). We now show that $\mu << \mathcal{L}^n$. Let $A \subseteq \mathbb{R}^n$ be bounded with $\mathcal{L}^n(A) = 0$. Fix $\epsilon > 0$ and choose an open bounded set $V \supseteq A$ such that $\mathcal{L}^n(V) < \epsilon$. Then by Theorem (1.1.1) and Hölder's inequality,

$$\begin{aligned}
\mu(A) \leq \mu(V) &\leq \liminf_{j \to +\infty} \mu_{k_j}(V) = \liminf_{j \to +\infty} \int_V f_{k_j} \, d\mathcal{L}^n \\
&\leq \liminf_{j \to +\infty} \|f_{k_j}\| L^p(V) \cdot \mathcal{L}^n(V)^{\frac{p-1}{p}} \\
&\leq C\epsilon^{\frac{p-1}{p}}.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary and $\frac{p-1}{p} > 0$, $\mu(A) = 0$, as required. Therefore $\mu << \mathcal{L}^n$.

(iv). By the Radon–Nikodym Theorem, there exists $f \in L^1_{loc}(\mathbb{R}^n)$ such that

$$\mu(A) = \int_A f \, d\mathcal{L}^n$$

for every Borel set $A \subseteq \mathbb{R}^n$.

(v). We prove that $f \in L^p(\mathbb{R}^n)$. Let $\phi \in \mathcal{C}_c(\mathbb{R}^n)$. Then

$$\begin{aligned}
\int_{\mathbb{R}^n} f\phi \, d\mathcal{L}^n = \int_{\mathbb{R}^n} \phi \, d\mu &= \lim_{j \to +\infty} \int_{\mathbb{R}^n} \phi \, d\mu_{k_j} \\
&= \lim_{j \to +\infty} \int_{\mathbb{R}^n} \phi f_{k_j} d\mathcal{L}^n \\
&\leq \sup_{k \in \mathbb{N}} \|f_{k_j}\|_{L^p}(\mathbb{R}^n) \|\phi\|_{L^q(\mathbb{R}^n)} \\
&\leq C\|\phi\|_{L^q(\mathbb{R}^n)}.
\end{aligned}$$

Thus

$$\|f\|_{L^p(\mathbb{R}^n)} = \sup_{\substack{\phi \in \mathcal{C}_c(\mathbb{R}^n) \\ \|\phi\|_{L^q(\mathbb{R}^n)=1}}} \left| \int_{\mathbb{R}^n} f\phi \, d\mathcal{L}^n \right| \leq C < +\infty,$$

and we see that $f \in L^p(\mathbb{R}^n)$.

(vi). Finally, we show that $f_{k_j} \rightharpoonup f$ in $L^p(\mathbb{R}^n)$. Fix $\epsilon > 0$. By the above,

$$\int_{\mathbb{R}^n} f_{k_j}\phi \, d\mathcal{L}^n \to \int_{\mathbb{R}^n} f\phi \, d\mathcal{L}^n$$

as $j \to +\infty$ for all $\phi \in \mathcal{C}_c(\mathbb{R}^n)$. Thus we may choose $J \in \mathbb{N}$ so large so that for all $j > J$,

$$\left| \int_{\mathbb{R}^n} f_{k_j}\phi - f\phi \, d\mathcal{L}^n \right| < \epsilon \tag{1.1.3}$$

{eq:1.9-3

for all $\phi \in \mathcal{C}_c(\mathbb{R}^n)$. Given $g \in L^q(\mathbb{R}^n)$, choose by the density of $\mathcal{C}_c(\mathbb{R}^n)$ in $L^q(\mathbb{R}^n)$ a function $\phi \in \mathcal{C}_c(\mathbb{R}^n)$ such that

$$\|g - \phi\|_{L^q(\mathbb{R}^n)} < \epsilon.$$

eq:1.9-3

Then by (1.1.3), Hölder's inequality, and the Principle of Uniform Boundedness, we have for all $j > J$

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} f_{k_j} g \, d\mathcal{L}^n - \int_{\mathbb{R}^n} fg \, d\mathcal{L}^n \right| &\leq \int_{\mathbb{R}^n} |f_{k_j} g - f_{k_j}\phi| \, d\mathcal{L}^n + \left| \int_{\mathbb{R}^n} f_{k_j}\phi - f\phi \, d\mathcal{L}^n \right| + \\
&\quad \int_{\mathbb{R}^n} |f\phi - fg| \, d\mathcal{L}^n \\
&\leq \epsilon + \int_{\mathbb{R}^n} |f_{k_j}||g - \phi| \, d\mathcal{L}^n + \int_{\mathbb{R}^n} |f||\phi - g| \, d\mathcal{L}^n \\
&\leq \epsilon + \epsilon\|f_{k_j}\|_{L^p(\mathbb{R}^n)} + \epsilon\|f\|_{L^p(\mathbb{R}^n)} \\
&\leq (2C+1)\epsilon.
\end{aligned}$$

The proof is complete. □

2. Hausdorff Measure

2.1. Definitions and Elementary Properties; Hausdorff Dimension.

Definition 2.1.1 ($\mathcal{H}^s_\delta$)**.** *Let $A \subseteq \mathbb{R}^n$, $0 \leq s < +\infty$, $0 < \delta \leq +\infty$. We define*

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

where

$$\alpha(s) := \frac{\pi^{\frac{s}{2}}}{\Gamma(1+\frac{s}{2})}$$

denotes the volume of the unit ball in $\mathbb{R}^s$.

Note in the above definition that s need not be an integer.

Definition 2.1.2 ($\mathcal{H}^s$, $s-$Dimensional Hausdorff Measure)**.** *Let $A \subseteq \mathbb{R}^n$, $0 \leq s < +\infty$. We define the $s-$dimensional Hausdorff measure $\mathcal{H}^s$ on $\mathbb{R}^n$ by*

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

Note that taking the limit as $\delta \to 0$ coincides with taking the supremum over $\delta > 0$, for, as $\delta \to 0$, we are taking the infimum over smaller and smaller sets. That is, if $\delta_1 < \delta_2$, then there exist coverings $\{C_j\}_{j=1}^{+\infty}$ of A such that $\operatorname{diam} C_j \leq \delta_2$ but $\operatorname{diam} C_j > \delta_1$.

Remark.

(i) Requiring $\delta \to 0$ forces the coverings to "follow the local geometry" of the set A;

(ii) Recall that

$$\mathcal{L}^n(B(x,r)) = \alpha(n) r^n$$

for all balls $B(x,r) \subseteq \mathbb{R}^n$. In fact if $s = k$ is an integer, then $\mathcal{H}^k$ coincides with the ordinary "$k-$dimensional surface area" on nice sets. This is the reason that the normalizing constant $\alpha(s)$ is included in the definition of $\mathcal{H}^s_\delta$.

t2.1-1 **Theorem 2.1.1.** *$\mathcal{H}^s$ is a Borel regular measure, $0 \leq s < +\infty$.*

Remark.

(i) Recall that this means that $\mathcal{H}^s$ is Borel and for each $A \subseteq \mathbb{R}^n$ there exists a Borel set B such that $A \subseteq B$ and $\mathcal{H}^s(A) = \mathcal{H}^s(B)$.

(ii) $\mathcal{H}^s$ is **not** *a Radon measure if $0 \leq s < n$, since $\mathbb{R}^n$ is not $\sigma-$finite with respect to $\mathcal{H}^s$.*

Proof.

(i). $\mathcal{H}^s_\delta$ is a measure. Choose $\{A_k\}_{k=1}^{+\infty} \subseteq \mathbb{R}^n$ and suppose that $A_k \subseteq \cup_{j=1}^{+\infty} C_j^k$, where $\operatorname{diam} C_j^k \leq \delta$. Then $\{C_j^k\}_{j,k=1}^{+\infty}$ covers $\cup_{k=1}^{+\infty} A_k$. Thus

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

Taking infima over all such covers $\{C_j^k\}_{k=1}^{+\infty}$ of A_k, we find

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{+\infty} A_k\right) \leq \sum_{k=1}^{+\infty} \mathcal{H}^s_\delta(A_k),$$

as required.

(ii). $\mathcal{H}^s$ is a measure. Choose $\{A_k\}_{k=1}^{+\infty} \subseteq \mathbb{R}^n$. Since $\mathcal{H}^s(\cup_{k=1}^{+\infty} A_k) = \sup_{\delta>0} \mathcal{H}^s_\delta(\cup_{k=1}^{+\infty} A_k)$, we have

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{+\infty} A_k\right) \leq \sum_{k=1}^{+\infty} \mathcal{H}^s_\delta(A_k) \leq \sum_{k=1}^{+\infty} \mathcal{H}^s(A_k).$$

Taking the limit as $\delta \to 0$ on the LHS shows that

$$\mathcal{H}^s\left(\bigcup_{k=1}^{+\infty} A_k\right) \leq \sum_{k=1}^{+\infty} \mathcal{H}^s(A_k).$$

(iii). $\mathcal{H}^s$ is a Borel measure. Choose $A, B \subseteq \mathbb{R}^n$ with $\operatorname{dist}(A,B) > 0$. Select $0 < \delta < \frac{1}{4}\operatorname{dist}(A,B)$. Let $A \cup B \subseteq \cup_{k=1}^{+\infty} C_k$ with $\operatorname{diam} C_k \leq \delta$.

Put

$$\mathcal{A} := \{C_j : C_j \cap A \neq \emptyset\}$$

and

$$\mathcal{B} := \{C_j : C_j \cap B \neq \emptyset\}.$$

Then $A \subseteq \cup_{C_j \in \mathcal{A}} C_j$ and $B \subseteq \cup_{C_j \in \mathcal{B}} C_j$, with $C_i \cap C_j = \emptyset$ if $C_i \in \mathcal{A}$, $C_j \in \mathcal{B}$. Thus

$$\begin{aligned}\sum -j = 1^{+\infty} \frac{\alpha(s)}{2^s}(\operatorname{diam} C_j)^s &\geq \sum_{C_j \in \mathcal{A}} \frac{\alpha(s)}{2^s}(\operatorname{diam} C_j)^s + \sum_{C_j \in \mathcal{B}} \frac{\alpha(s)}{2^s}(\operatorname{diam} C_j)^s \\ &\geq \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B).\end{aligned}$$

Taking the infimum over all such sets $\{C_j\}_{j=1}^{+\infty}$, $0 < \delta < \frac{1}{4}\operatorname{dist}(A,B)$, we find

$$\mathcal{H}^s_\delta(A \cup B) \geq \mathcal{H}^s_\delta(A) + \mathcal{H}^s_\delta(B).$$

Letting $\delta \to 0$, we obtain

$$\mathcal{H}^s(A \cup B) \geq \mathcal{H}^s(A) + \mathcal{H}^s(B).$$

Consequently

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

for all $A, B \subseteq \mathbb{R}^n$ with $\operatorname{dist}(A,B) > 0$. By Caratheodory's Criterion, $\mathcal{H}^s$ is a Borel measure.

(iv). $\mathcal{H}^s$ is Borel regular. First note that $\operatorname{diam} \overline{C} = \operatorname{diam} C$ for all $C \subseteq \mathbb{R}^n$. Thus

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

Choose $A \subseteq \mathbb{R}^n$ such that $\mathcal{H}^s(A) < +\infty$. Then $\mathcal{H}^s_\delta(A) < +\infty$ for all $\delta > 0$. For each $k \geq 1$, choose closed sets $\{C_j^k\}_{j=1}^{+\infty}$ so that $\operatorname{diam} C_j^k \leq \frac{1}{k}$, $A \subseteq \cup_{j=1}^{+\infty} C_j^k$, and

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

Put $A_k := \cup_{j=1}^{+\infty} C_j^k$ and $B := \cap_{k=1}^{+\infty} A_k$. Then B is Borel. Also $A \subseteq A_k$ for each $k \in \mathbb{N}$, so $A \subseteq B$. Moreover, since $B \subseteq A_k$ for each k,

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

Letting $k \to +\infty$, we find

$$\mathcal{H}^s(B) \leq \mathcal{H}^s(A).$$

But since $A \subseteq B$, we have by monotonicity

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

The proof is complete. □

t2.1-2

Theorem 2.1.2 (Elementary Properties of Hausdorff Measure)**.**

(i) $\mathcal{H}^0$ *is counting measure;*
(ii) $\mathcal{H}^1 = \mathcal{L}^1$ *on* $\mathbb{R}$*;*
(iii) $\mathcal{H}^s \equiv 0$ *on* $\mathbb{R}^n$ *for all* $s > n$*;*
(iv) $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ *for all* $\lambda > 0$, $A \subseteq \mathbb{R}^n$*;*
(v) $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$ *for each affine isometry* $L : \mathbb{R}^n \to \mathbb{R}^n$, $A \subseteq \mathbb{R}^n$.

Proof.
(iv). Fix $0 < \delta \leq +\infty$, and suppose that $A \subseteq \cup_{j=1}^{+\infty} C_j$, with $\operatorname{diam} C_j \leq \delta$. Then $\lambda A \subseteq \cup_{j=1}^{+\infty} \lambda C_j$, and $\operatorname{diam} \lambda C_j = \lambda \operatorname{diam} C_j \leq \lambda\delta$. Thus

$$\begin{aligned}\lambda^s \sum_{j=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} C_j)^s &= \sum_{j=1}^{+\infty} \frac{\alpha(s)}{2^s} (\lambda \operatorname{diam} C_j)^s \\ &\geq \mathcal{H}^s_{\lambda\delta}(\lambda A).\end{aligned}$$

Taking the infimum over all such covers $\{C_j\}_{j=1}^{+\infty}$ of A, we deduce

$$\lambda^s \mathcal{H}^s_\delta(A) \geq \mathcal{H}^s_{\lambda\delta}(\lambda A),$$

and taking the limit as $\delta \to 0$ shows

$$\lambda^s \mathcal{H}^s(A) \geq \mathcal{H}^s(\lambda A.)$$

The reverse inequality may be shown similarly.
(v). This follows at once from (iv) along with the translation invariance of $\mathcal{H}^s$.
(i). First note that $\alpha(0) = 1$. Thus obviously $\mathcal{H}^0(\{a\}) = 1$ for all $a \in \mathbb{R}^n$, and (i) follows.
(ii). Choose $A \subseteq \mathbb{R}$ and $\delta > 0$. Then

$$\begin{aligned}\mathcal{L}^1(A) &= \inf\left\{\sum_{j=1}^{+\infty} \operatorname{diam} C_j : A \subseteq \bigcup_{j=1}^{+\infty} C_j\right\} \\ &\leq \inf\left\{\sum_{j=1}^{+\infty} \operatorname{diam} C_j : A \subseteq \bigcup_{j=1}^{+\infty} C_j, \operatorname{diam} C_j \leq \delta\right\} \\ &= \mathcal{H}^1_\delta(A) \\ &\leq \mathcal{H}^1(A).\end{aligned}$$

On the other hand, set $I_k := [k\delta, (k+1)\delta]$, $k \in \mathbb{Z}$. Then $\operatorname{diam}(C_j \cap I_k) \leq \delta$, and, since $\cup_{k=1}^{+\infty} C_j \cap I_k = C_j$,

$$\sum_{k=-\infty}^{+\infty} \operatorname{diam}(C_j \cap I_k) \leq \operatorname{diam} C_j.$$

Hence,

$$\begin{aligned}\mathcal{L}^1(A) &= \inf\left\{\sum_{j=1}^{+\infty} \operatorname{diam} C_j : A \subseteq \bigcup_{j=1}^{+\infty} C_j\right\} \\ &\geq \inf\left\{\sum_{j=1}^{+\infty}\sum_{k=-\infty}^{+\infty} \operatorname{diam}(C_j \cap I_k) : A \subseteq \bigcup_{j=1}^{+\infty} C_j\right\} \\ &= \mathcal{H}^1_\delta(A).\end{aligned}$$

Therefore $\mathcal{L}^1 = \mathcal{H}^1_\delta$ for all $\delta > 0$, so that taking the supremum over all $\delta > 0$, we have $\mathcal{L}^1 = \mathcal{H}^1$ on $\mathbb{R}$.

(iii). Fix an integer $m \geq 1$. The unit cube $Q(n)$ in $\mathbb{R}^n$ may be decomposed into m^n cubes with side length $\frac{1}{m}$ and diameter $\frac{\sqrt{n}}{m}$. Thus

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

and the RHS tends to zero as $m \to +\infty$ if $s > n$. Hence $\mathcal{H}^s(Q(n)) = 0$, so $\mathcal{H}^s \equiv 0$. The proof is complete. □

A convenient way to check that $\mathcal{H}^s$ vanishes on a set $A \subseteq \mathbb{R}^n$ is the following lemma.

l2-1-1 **Lemma 2.1.1.** *If $A \subseteq \mathbb{R}^n$ and $\mathcal{H}^s_\delta(A) = 0$ for some $0 < \delta \leq +\infty$, then $\mathcal{H}^s(A) = 0$.*

Proof. The conclusion is obvious if $s = 0$, and so we may assume that $s > 0$.

Fix $\epsilon > 0$. There exist sets $\{C_j\}_{j=1}^{+\infty}$ such that $A \subseteq \cup_{j=1}^{+\infty} C_j$ and

$$\sum_{j=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} C_j)^s \leq \epsilon.$$

In particular for each $j \in \mathbb{N}$,

$$\operatorname{diam} C_j \leq 2\left(\frac{\epsilon}{\alpha(s)}\right)^{\frac{1}{s}} =: \delta(\epsilon).$$

Hence $\mathcal{H}^s_{\delta(\epsilon)} < \epsilon$. But since $\delta(\epsilon) \to 0$ and $\epsilon \to 0$, we have

$$\mathcal{H}^s(A) = 0.$$

The proof is complete. □

We next want to define the *Hausdorff dimension* of a subset of $\mathbb{R}^n$.

l2.1-2 **Lemma 2.1.2.** *Let $A \subseteq \mathbb{R}^n$ and $0 \leq s < t < +\infty$.*

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

Proof.

(i). Let $\mathcal{H}^s(A) < +\infty$ and $\delta > 0$. Then there exist sets $\{C_j\}_{j=1}^{+\infty}$ such that $A \subseteq \cup_{j=1}^{+\infty} C_j$, $\operatorname{diam} C_j \leq \delta$, and

$$\sum_{j=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} C_j)^s \leq \mathcal{H}^s_\delta(A) + 1 \leq \mathcal{H}^s(A) + 1.$$

Then

$$\begin{aligned}
\mathcal{H}^t_\delta(A) &\leq \sum_{j=1}^{+\infty} \frac{\alpha(t)}{2^t}(\operatorname{diam} C_j)^t \\
&= \frac{\alpha(t)}{\alpha(s)} 2^{s-t} \sum_{j=1}^{+\infty} \frac{\alpha(s)}{2^s}(\operatorname{diam} C_j)^s \cdot (\operatorname{diam} C_j)^{t-s} \\
&\leq \frac{\alpha(t)}{\alpha(s)} 2^{s-t} \delta^{t-s} (\mathcal{H}^s(A) + 1).
\end{aligned}$$

Sending $\delta \to 0$, we conclude that $\mathcal{H}^t(A) = 0$. This proves (i).

(ii). Assertion (ii) follows at once from (i), by contrapositive. The proof is complete. □

Definition 2.1.3 (Hausdorff Dimension)**.** *We define the Hausdorff dimension of a set $A \subseteq \mathbb{R}^n$ by*

$$\mathcal{H}_{\dim}(A) := \inf\{0 \leq s < +\infty : \mathcal{H}^s(A) = 0.\}$$

Remark. *Observe for any set $A \subseteq \mathbb{R}^n$ that $\mathcal{H}_{\dim}(A) \leq n$. Let $s := \mathcal{H}_{\dim}(A)$. Then by the preceding lemma, $\mathcal{H}^t(A) = 0$ for all $t > s$ and $\mathcal{H}^t(A) = +\infty$ for all $t < s$. Moreover, $\mathcal{H}^s(A)$ may be any number between 0 and $+\infty$, inclusive. The point is that $s = \mathcal{H}_{\dim}$ is the only number such that $\mathcal{H}^s(A)$ can be a positive finite number for any $A \subseteq \mathbb{R}^n$.*

Also note that $\mathcal{H}_{\dim}(A)$ need not be an integer. Even if $\mathcal{H}_{\dim}(A) = k$ is an integer and $0 < \mathcal{H}^k(A) < +\infty$, A need not be a "$k-$dimensional surface" in any sense, and may be extremely complicated geometrically. Examples include Cantor–like *subsets A of $\mathbb{R}^n$ and other fractals.*

2.2. **Isodiametric Inequality; $\mathcal{H}^n = \mathcal{L}^n$.**

We want to prove that $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$, where $n \in \mathbb{N}$. Recall that $\mathcal{L}^n$ is defined as the $n-$fold product of one–dimensional Lebesgue measure $\mathcal{L}^1$, so that

$$\mathcal{L}^1(A) := \inf\left\{\sum_{i=1}^{n} \mathcal{L}^n(Q_i) : Q_i \text{ cubes }, A \subseteq \bigcup_{i=1}^{n} Q_i\right\}.$$

On the other hand, $\mathcal{H}^n$ is computed in terms of arbitrary coverings of small diameter.

12.2-1 **Lemma 2.2.1.** *Let $f : \mathbb{R}^n \to [0, +\infty]$ be L^n-measurable. Then the region "under the graph" of f,*

$$A := \{(x, y) : x \in \mathbb{R}^n, y \in \mathbb{R}, 0 \leq y \leq f(x)\}$$

is $\mathcal{L}^{n+1}-$measurable.

Proof. Define

$$B := \{x \in \mathbb{R}^n : f(x) = +\infty\}$$

and

$$C := \{x \in \mathbb{R}^n : 0 \leq f(x) < +\infty.\}$$

Also define

$$C_{j,k} := \left\{x \in C : \frac{j}{k} \leq f(x) < \frac{j+1}{k}\right\}, \quad j \in \mathbb{N}_0, \quad k \in \mathbb{N},$$

so that $C = \cup_{j=0}^{+\infty} C_{j,k}$. Finally, put

$$D_k := \bigcup_{j=0}^{+\infty} \left(C_{j,k} \times \left[0, \frac{j}{k}\right]\right) \cup (B \times [0, +\infty]),$$

$$E_k := \bigcup_{j=0}^{+\infty} \left(C_{j,k} \times \left[0, \frac{j+1}{k}\right]\right) \cup (B \times [0, +\infty]).$$

Clearly D_k and E_k are $\mathcal{L}^{n+1}$ measurable, and we have for each $k \in \mathbb{N}$ $D_k \subseteq A \subseteq E_k$. Write $D := \cup_{k=1}^{+\infty} D_k$ and $E := \cap_{k=1}^{+\infty} E_k$. Then also $D \subseteq A \subseteq E$, with D and E both $\mathcal{L}^{n+1}-$measurable. Now for any $\mathcal{L}^{n+1}-$measurable set F with $\mathcal{L}^{n+1}(F) < +\infty$,

$$\mathcal{L}^{n+1}((E \setminus D) \cap F) \leq \mathcal{L}^{n+1}((E_k \setminus D_k) \cap F) \leq \frac{1}{k}\mathcal{L}^n(F),$$

and the RHS tends to zero as $k \to +\infty$. Thus $\mathcal{L}^{n+1}((E \setminus D) \cap F) = 0$, and, because F was arbitrary, $\mathcal{L}^{n+1}(E \setminus D) = 0$. Hence $\mathcal{L}^{n+1}(A \setminus D) = 0$, and consequently A is $\mathcal{L}^{n+1}-$measurable. □

We now define the process of Steiner symmetrization, which takes a bounded Borel–measurable set $A \subseteq \mathbb{R}^n$ and transforms A into a set $\widetilde{A}$ having the same Lebesgue measure such that $\operatorname{diam}(\widetilde{A}) \leq \operatorname{diam}(A)$.

Fix $a, b \in \mathbb{R}^n$, $\|a\| = 1$. We define

$$L_b^a := \{b + ta : t \in \mathbb{R}\}, \text{ the line through } b \text{ in the direction of } a,$$

and

$$P_a := \{x \in \mathbb{R}^n : x \cdot a = 0\}, \text{ the plane through the origin perpendicular to } a.$$

Definition 2.2.1 (Steiner Symmetrization)**.** *Choose $a \in \mathbb{R}^n$ with $\|a\| = 1$, and let $A \subseteq \mathbb{R}^n$. We define the Steiner symmetrization of A with respect to the hyperplane P_a to be the set*

$$S_a(A) := \bigcup_{\substack{b \in P_a \\ A \cap L_b^a \neq \emptyset}} \left\{ b + ta : \|t\| \leq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a)\right\}.$$

Note that the Steiner symmetrization is the union of all line segments $b + ta$ of length less than $\mathcal{H}^1(A \cap L_b^a)$, where b is in the plane through the origin perpendicular to a and there exists $x \in A$ such that $b + ta = x$.

l2.2-2 **Lemma 2.2.2** (Properties of Steiner Symmetrization)**.**

(i) $\operatorname{diam} S_a(A) \leq \operatorname{diam} A$.

(ii) If A is $\mathcal{L}^n-$measurable, then so is $S_a(A)$, and $\mathcal{L}^n(S_a(A)) = \mathcal{L}^n(A)$.

Proof.

(i). Statement (i) is trivial if $\operatorname{diam} A = +\infty$, so we may assume that $\operatorname{diam} A < +\infty$. We may also suppose that A is closed, for

$$\operatorname{diam} A^\circ = \operatorname{diam} A = \operatorname{diam} \overline{A}.$$

Fix $\epsilon > 0$ and choose $x, y \in S_a(A)$ such that

$$\operatorname{diam} S_a(A) \leq \|x - y\| + \epsilon.$$

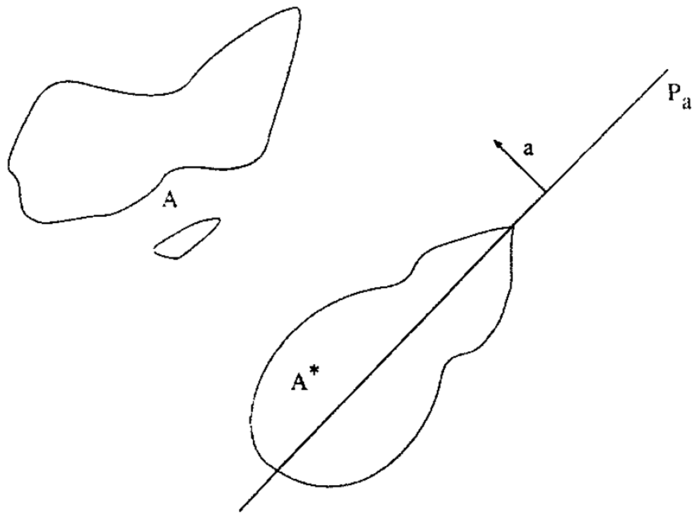

FIGURE 2.2.1. Steiner Symmetrization.

Write $b := x - (x \cdot a)a$ and $c := y - (y \cdot a)a$. Then $b, c \in P_a$. Put

$$r := \inf\{t : b + ta \in A\},$$
$$s := \sup\{t : b + ta \in A\},$$
$$u := \inf\{t : c + ta \in A\},$$
$$v := \sup\{t : c + ta \in A\}.$$

Without loss of generality, we may assume that $v - r \geq s - u$. Then

$$\begin{aligned} v - r &\geq \frac{1}{2}(v - r) + \frac{1}{2}(s - u) \\ &= \frac{1}{2}(s - r) + \frac{1}{2}(v - u) \\ &\geq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a) + \frac{1}{2}\mathcal{H}^1(A \cap L_c^a). \end{aligned}$$

Now, $|x \cdot a| \leq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a)$, $|y \cdot a| \leq \frac{1}{2}\mathcal{H}^1(A \cap L_b^a)$, and consequently,

$$v - r \geq |x \cdot a| + |y \cdot a| \geq |x \cdot a - y \cdot a|.$$

Hence,

$$\begin{aligned} (\operatorname{diam} S_a(A) - \epsilon)^2 &\leq \|x - y\|^2 \\ &= \|x\|^2 - 2x \cdot y + \|y\|^2 \\ &= \|b\|^2 + 2(x \cdot a)(b \cdot a) + |x\dot{a}|^2 - 2(b + (x \cdot a)a) \cdot (c + (y \cdot a)a) + \|c\|^2 + \\ &\qquad 2(y \cdot a)(b \cdot a) + |y \cdot a|^2 \\ &= (\|b\|^2 - 2b \cdot c + \|c\|^2) + (|x \cdot a|^2 - 2(x \cdot a)(y \cdot a) + |y \cdot a|^2) + \\ &\qquad 2(x \cdot a)(b \cdot a) - 2(b \cdot a)(y \cdot a) - 2(c \cdot a)(x \cdot a) + 2(y \cdot a)(b \cdot a) \\ &= \|b - c\|^2 + \|x \cdot a - y \cdot a\|^2 \end{aligned}$$

$$\begin{aligned}
&\leq \|b-c\|^2 + (v-r)^2 \\
&= \|b\|^2 - 2b\cdot c + \|c\|^2 + v^2 - 2rv + r^2 \\
&= (\|b\|^2 + 2b\cdot ra + \|ra\|^2) - 2(b\cdot c - b\cdot va - c\cdot ra - rv\|a\|^2) + \\
&\qquad (\|c\|^2 + 2c\cdot va + \|va\|^2) \\
&= \|(b+ra) - (c+va)\|^2 \\
&\leq (\operatorname{diam} A)^2,
\end{aligned}$$

since $b, c \perp a$ and A is closed, so that $b + ra, c + va \in A$. Thus $\operatorname{diam} S_a(A) - \epsilon \leq \operatorname{diam} A$, and since $\epsilon > 0$ was arbitrary, this proves (i).

(ii). Since $\mathcal{L}^n$ is rotation invariant, we may assume that $a = e_n$. Then $P_a = P_{e_n} = \mathbb{R}^{n-1}$. Since $\mathcal{L}^1 = \mathcal{H}^1$ on $\mathbb{R}$, Tonelli's Theorem implies that the map $f : \mathbb{R}^{n-1} \to \mathbb{R}$ defined by $f(b) = \mathcal{H}^1(A \cap L_b^a)$ is $\mathcal{L}^{n-1}-$measurable and $\mathcal{L}^n(A) = \int_{\mathbb{R}^{n-1}} f(b)\, d\mathcal{L}^{n-1}(b)$, for

$$\int_{\mathbb{R}^{n-1}} f(b)\, d\mathcal{L}^{n-1}(b) = \int_{\mathbb{R}^{n-1}} \mathcal{L}^1(A \cap L_b^a)\, d\mathcal{L}^{n-1}(b) = \mathcal{L}^n(A).$$

Therefore

$$S_a(A) = \left\{(b, y) : 0 \leq |y| \leq \frac{f(b)}{2}\right\} \setminus \{(b, 0) : L_b^a \cap A = \emptyset\}$$

is $\mathcal{L}^n-$measurable by Lemma (2.2.1), and

$$\begin{aligned}
\mathcal{L}^n(S_a(A)) &= \int_{\mathbb{R}}^n \mathbb{1}_{S_a(A)}\, d\mathcal{L}^n = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \mathbb{1}_{S_a(A)}\, d\mathcal{L}^1\, d\mathcal{L}^{n-1} \\
&= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} (\mathbb{1}_{S_a(A)})_{(e_1,\ldots,e_{n-1})}(y)\, d\mathcal{L}^1(y)\, d\mathcal{L}^{n-1} \\
&= \int_{\mathbb{R}^{n-1}} \int_{-f(b)/2}^{f(b)/2} d\mathcal{L}^1\, d\mathcal{L}^{n-1} \\
&= \int_{\mathbb{R}^{n-1}} f(b)\, d\mathcal{L}^{n-1}(b) = \mathcal{L}^n(A).
\end{aligned}$$

The proof is complete. □

Remark. *In proving $\mathcal{H}^n = \mathcal{L}^n$ below, notice that we use only statement (ii) above in the special case that a is a standard coordinate vector. Since $\mathcal{H}^n$ is obviously rotation invariant, we in fact prove that $\mathcal{L}^n$ is rotation invariant also.*

t2.2-1 **Theorem 2.2.1** (Isodiametric Inequality)**.** *For all sets $A \subseteq \mathbb{R}^n$,*

$$\mathcal{L}^n(A) \leq \frac{\alpha(n)}{2^n}(\operatorname{diam} A)^n.$$

Remark.

(i) Geometrically, the isodiametric inequality says that of all sets of fixed diameter in $\mathbb{R}^n$, the $n-$sphere has greatest volume.

(ii) This inequality is particularly interesting because it is not necessarily the case that A is contained in a ball of diameter $\operatorname{diam} A$, *for in $\mathbb{R}^2$ consider the case of an equilateral triangle*

with side length 1. The smallest closed ball B which inscribes the triangle has radius $1/\sqrt{3}$, so

$$\operatorname{diam} B = \frac{2}{\sqrt{3}} > 1.$$

Proof. If $\operatorname{diam} A = +\infty$, the inequality is trivial. Therefore we may assume that $\operatorname{diam} A < +\infty$.

Let $\{e_1, \dots, e_n\}$ be the standard basis for $\mathbb{R}^n$. Define $A_1 := S_{e_1}(A)$, $A_2 := S_{e_2}(A_1), \dots, A_n := S_{e_n}(A_{n-1})$. Write $A^* := A_n$.

(i). We first show that A^* is symmetric with respect to the origin. We use induction. Clearly A_1 is symmetric with respect to P_{e_1}. Let k be an integer such that $1 \le k < n$ and suppose that A_k is symmetric with respect to $P_{e_1}, \dots, P_{e_k}$. Clearly $A_{k+1} = S_{e_{k+1}}(A_k)$ is symmetric with respect to $P_{e_{k+1}}$. Fix $1 \le j < k$ and let $S_j : \mathbb{R}^n \to \mathbb{R}^n$ be the reflection through P_{e_j}. Let $b \in P_{e_{k+1}}$. Since A_k is symmetric with respect to $P_{e_1}, \dots, P_{e_k}$ by the induction hypothesis and $1 \le j \le k$, we have $S_j(A_k) = A_k$, and so

$$\mathcal{H}^1(A_k \cap L_b^{e_{k+1}}) = \mathcal{H}^1(A_k \cap L_{S_j b}^{e_{k+1}}).$$

Consequently

$$\{t \in \mathbb{R} : b + te_{k+1} \int A_{k+1}\} = \{t \in \mathbb{R} : S_j b + te_{k+1} \in A_{k+1}\}.$$

Thus $S_j(A_{k+1}) = A_{k+1}$, that is, A_{k+1} is symmetric with respect to P_{e_j}. Since j was arbitrary, $A^* = A_n$ is symmetric with respect to $P_{e_1}, \dots, P_{e_n}$, and so with respect to the origin.

(ii). We show that

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

Choose $x \in A^*$. Then $-x \in A^*$ by (i), and so $\operatorname{diam} A^* \ge 2|x|$. Thus $A^* \subseteq B(0, \frac{1}{2}\operatorname{diam} A^*)$, and it follows by monotonicity of the Lebesgue measure

$$\mathcal{L}^n(A^*) \le \mathcal{L}^n\left(B\left(0, \frac{1}{2}\operatorname{diam} A^*\right)\right) = \frac{\alpha(n)}{2^n}(\operatorname{diam} A^*)^2.$$

(iii). We now prove the isodiametric inequality. Note that $\overline{A}$ is $\mathcal{L}^n-$measurable, and thus the above Lemma (2.2.2) implies that

$$\mathcal{L}^n((\overline{A})^*) = \mathcal{L}^n(\overline{A}),$$

as well as

$$\operatorname{diam}(\overline{A})^* \le \operatorname{diam} \overline{A}.$$

Hence, monotonicity of the Lebesgue measure together with (ii) give

$$\begin{aligned}
\mathcal{L}^n(A) \le \mathcal{L}^n(\overline{A}) &= \mathcal{L}^n((\overline{A})^*) \\
&\le \frac{\alpha(n)}{2^n}(\operatorname{diam}(\overline{A})^*)^n \\
&\le \frac{\alpha(n)}{2^n}(\operatorname{diam}(\overline{A}))^n \\
&= \frac{\alpha(n)}{2^n}(\operatorname{diam} A)^n.
\end{aligned}$$

The proof is complete. □

t2.2-2 **Theorem 2.2.2.** *On* $\mathbb{R}^n$, $\mathcal{L}^n = \mathcal{H}^n$.

Proof. (i). We first show that $\mathcal{L}^n(A) \leq \mathcal{H}^n(A)$ for all $A \subseteq \mathbb{R}^n$. Fix $\delta > 0$. Choose sets $\{C_j\}_{j=1}^{+\infty}$ such that $A \subseteq \cup_{j=1}^{+\infty} C_j$ and $\operatorname{diam} C_j \leq \delta$. Then by monotonicity and the Isodiametric Inequality (cf. (2.2.1)),

$$\mathcal{L}^n(A) \leq \sum_{j=1}^{+\infty} \mathcal{L}^n(C_j) \leq \sum_{j=1}^{+\infty} \frac{\alpha(n)}{2^n} (\operatorname{diam} C_j)^n.$$

Taking the infimum of the RHS over all cover countable covers of A with diameter less than δ, we obtain $\mathcal{L}^n(A) \leq H^n_\delta(A)$. Taking the limit as $\delta \to 0$, we have

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

as required.

(ii). From the definition of $\mathcal{L}^n$ as the $n-$fold product of $\mathcal{L}^1 \times \cdots \times \mathcal{L}^1$, we see that for all $A \subseteq \mathbb{R}^n$ and $\delta > 0$,

$$\mathcal{L}^n(A) = \inf \left\{ \sum_{i=1}^{+\infty} \mathcal{L}^n(Q_i) : Q_i \text{ cubes, } A \subseteq \bigcup_{i=1}^{+\infty}, \operatorname{diam} Q_i \leq \delta \right\}.$$

We may consider only cubes parallel to the coordinate axes in $\mathcal{L}^n$.

(iii). We now show that $\mathcal{H}^n$ is absolutely continuous with respect to $\mathcal{L}^n$. Set $C_n := \frac{\alpha(n)}{2^n}$. Then for each cube $Q \subseteq \mathbb{R}^n$,

$$\frac{\alpha(n)}{2^n} (\operatorname{diam} Q)^n = C_n \mathcal{L}^n(Q).$$

Thus for any $A \subseteq \mathbb{R}^n$,

$$\begin{aligned}
\mathcal{H}^n_\delta(A) &= \inf \left\{ \sum_{i=1}^{n} \frac{\alpha(n)}{2^n} (\operatorname{diam} U_i)^n : A \subseteq \bigcup_{i=1}^{+\infty} U_i, \operatorname{diam} U_i \leq \delta \right\} \\
&\leq \inf \left\{ \sum_{i=1}^{+\infty} \frac{\alpha(n)}{2^n} (\operatorname{diam} Q_i)^n : Q_i \text{ cubes }, A \subseteq \bigcup_{i=1}^{+\infty} Q_i, \operatorname{diam} Q_i \leq \delta \right\} \\
&= C_n \mathcal{L}^n(A).
\end{aligned}$$

Taking the supremum over all $\delta > 0$, we've:

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

Thus $\mathcal{H}^n(A) = 0$ whenever $\mathcal{L}^n(A) = 0$. This proves (iii).

(iv). We now show that $\mathcal{H}^n(A) \leq \mathcal{L}^n(A)$ for all $A \subseteq \mathbb{R}^n$. To this end, fix $\delta > 0$ and $\epsilon > 0$. We may choose cubes $\{Q_i\}_{i=1}^{+\infty} \subseteq \mathbb{R}^n$ such that $A \subseteq \cup_{i=1}^{+\infty} Q_i$, $\operatorname{diam} Q_i \leq \delta$, and

$$\sum_{i=1}^{+\infty} \mathcal{L}^n(Q_i) < \mathcal{L}^n(A) + \epsilon.$$

Now for each $i \in \mathbb{N}$ there exist disjoint closed balls $\{B^i_k\}_{k=1}^{+\infty} \subseteq Q_i^\circ$ such that

$$\operatorname{diam} B^i_k \leq \delta$$

and

$$\mathcal{L}^n\left(Q_i \setminus \bigcup_{k=1}^{+\infty} B_k^i\right) = \mathcal{L}^n\left(Q_i^\circ \setminus \bigcup_{k=1}^{+\infty} B_k^i\right) = 0.$$

Since $\mathcal{H}^n, \mathcal{H}_\delta^n$ are absolutely continuous with respect to $\mathcal{L}^n$ by (iii), $\mathcal{H}^n(Q_i \setminus \cup_{k=1}^{+\infty} B_k^i) = \mathcal{H}_\delta^n(Q_i \setminus \cup_{k=1}^{+\infty} B_k^i) = 0$. Therefore $\mathcal{H}^n(Q_i) = \mathcal{H}^n(\cup_{k=1}^{+\infty} B_k^i)$ and $\mathcal{H}_\delta^n(Q_i) = \mathcal{H}_\delta^n(\cup_{k=1}^{+\infty} B_k^i)$, and we have

$$\begin{aligned}
\mathcal{H}_\delta^n(A) &\le \sum_{i=1}^{+\infty} \mathcal{H}_\delta^n(Q_i) = \sum_{i=1}^{+\infty} \mathcal{H}_\delta^n\left(\bigcup_{k=1}^{+\infty} B_k^i\right) \le \sum_{i=1}^{+\infty}\sum_{k=1}^{+\infty} \mathcal{H}_\delta^n(B_k^i) \le \sum_{i=1}^{+\infty}\sum_{k=1}^{+\infty} \mathcal{H}^n(B_k^i) \\
&= \sum_{i=1}^{+\infty}\sum_{k=1}^{+\infty} \frac{\alpha(n)}{2^n}(\operatorname{diam} B_k^i)^n = \sum_{i=1}^{+\infty}\sum_{k=1}^{+\infty} \mathcal{L}^n(B_k^i) = \sum_{i=1}^{+\infty} \mathcal{L}^n\left(\bigcup_{k=1}^{\infty} B_k^i\right) \\
&= \sum_{i=1}^{+\infty}\sum_{i=1}^{+\infty} \mathcal{L}^n(Q_i) < \mathcal{L}^n(A) + \epsilon.
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, it follows $\mathcal{H}^n(A) \le \mathcal{L}^n(A)$. The proof is complete. □

2.3. **Densities.** We first recall the Lebesgue Density Theorem:

Theorem (Lebesgue Density Theorem). *Let $E \subseteq \mathbb{R}^n$ be a Lebesgue measurable set. For any $r > 0$ and $x \in \mathbb{R}^n$, define the approximate Lebesgue density of E in the $r-$neighborhood of x by*

$$d_r(x) := \frac{\mathcal{L}^n(B(x,r) \cap E)}{\alpha(n) r^n}.$$

Further define the Lebesgue density of E at x by

$$d(x) := \lim_{r\to 0} d_r(x).$$

Then

$$d(x) = \lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap E)}{\alpha(n) r^n} = \begin{cases} 1, & \text{for } \mathcal{L}^n - \text{a.e. } x \in E, \\ 0, & \text{for } \mathcal{L}^n - \text{a.e. } x \in \mathbb{R}^n \setminus E. \end{cases}$$

Since $\mathcal{H}^n = \mathcal{L}^n$ for $n \in \mathbb{N}$, the above result clearly holds for $\mathcal{H}^n$ as well. We want to develop some analogous results for lower–dimensional Hausdorff measures. Thus we assume throughout this section that $0 < s < n$.

We first establish a theorem that tells us the lower–dimensional Hausdorff density of a set at a.e. point outside the set is zero.

t2.3-1 **Theorem 2.3.1.** *Assume that $E \subseteq \mathbb{R}^n$ with E $\mathcal{H}^s-$measurable and $\mathcal{H}^s(E) < +\infty$. Then*

$$\lim_{r\to 0} \frac{\mathcal{H}^s(B(x,r)\cap E)}{\alpha(s) r^s} = 0$$

for $\mathcal{H}^s-$a.e. $x \in \mathbb{R}^n \setminus E$.

Proof. Fix $t > 0$ and define

$$A_t := \left\{ x \in \mathbb{R}^n \setminus E : \limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

It suffices to show that $\mathcal{H}^s(A_t) = 0$.

Note that $\mathcal{H}^s \llcorner E$ is a Radon measure, and so, if we fix $\epsilon > 0$, there exists a compact set $K \subseteq E$ such that

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

Set $U := \mathbb{R}^n \setminus K$. Then U is open and $A_t \subseteq U$ because $K \subseteq E$. Fix $\delta > 0$ and consider

$$\mathcal{F} := \left\{ B(x,r) : B(x,r) \subseteq U, 0 < r < \delta, \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

By the Vitali Covering Lemma, there exists a countable family of balls $\{B(x_i, r_i)\}_{i=1}^{+\infty}$ such that

$$A_t \subseteq \bigcup_{i=1}^{+\infty} B(x_i, 5r_i).$$

Thus by monotonicity

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(A_t) &\leq \mathcal{H}^s_{10\delta}\left(\bigcup_{i=1}^{+\infty} B(x_i, 5r_i) \right) \leq \sum_{i=1}^{+\infty} \frac{\alpha(s)}{2^s} (10 r_i)^s \leq \sum_{i=1}^{+\infty} 5^s \alpha(s) r^s \\
&\leq \frac{5^s}{t} \sum_{i=1}^{+\infty} \mathcal{H}^s(B(x_i, r_i) \cap E) \leq \frac{5^s}{t} \mathcal{H}^s(U \cap E) = \frac{5^s}{t} \mathcal{H}^s(E \setminus K) \\
&\leq \frac{5^s}{t} \epsilon.
\end{aligned}$$

Letting $\delta \to 0$, we obtain $\mathcal{H}^s(A_t) \leq \frac{5^s}{t}\epsilon$. Since $\epsilon > 0$ was arbitrary, we have $\mathcal{H}^s(A_t) = 0$ for each $t > 0$. The proof is complete. ☐

Now we prove that the lower–dimensional Hausdorff density of a set at a.e. point in the set is nonzero. Note that this contrasts with the Lebesgue Density Theorem: the density may not be 1. However, it is bounded below if we replace the limit with limit superior.

t2.3-2 **Theorem 2.3.2.** *Assume that $E \subseteq \mathbb{R}^n$ with $E \mathcal{H}^s-$measurable and $\mathcal{H}^s(E) < +\infty$. Then*

$$\frac{1}{2^s} \leq \limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} \leq 1$$

for $\mathcal{H}^s-$a.e. $x \in E$.

Remark. *It is possible to have*

$$\limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} < 1$$

and

$$\liminf_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} = 0$$

for $\mathcal{H}^s-$a.e. $x \in E$, even if $0 < \mathcal{H}^s(E) < +\infty$.

Proof. (i) We first show the upper inequality. Fix $\epsilon > 0$, $t > 1$, and define

$$B_t := \left\{ x \in E : \limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

Since $\mathcal{H}^s \llcorner E$ is Radon, there exists an open set U containing B_t such that

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

Define

$$\mathcal{F} := \left\{ B(x,r) : B(x,r) \subseteq U, 0 < r < \delta, \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} > t \right\}.$$

By a corollary of the Vitali Covering Lemma, there exists a countable family of disjoint balls $\{B(x_i, r_i)\}_{i=1}^{+\infty}$ such that

$$B_t \subseteq \left(\bigcup_{i=1}^{m} B(x_i, r_i) \right) \cup \left(\bigcup_{i=m+1}^{+\infty} B(x_i, 5r_i) \right).$$

Thus

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(B_t) &\leq \mathcal{H}^s_{10\delta}\left(\bigcup_{i=1}^{m} B(x_i, r_i) \right) + \mathcal{H}^s_{10\delta}\left(\bigcup_{i=m+1}^{+\infty} B(x_i, 5r_i) \right) \\
&\leq \sum_{i=1}^{m} \frac{\alpha(s)}{2^s}(2r_i)^s + \sum_{i=m+1}^{+\infty} \frac{\alpha(s)}{2^s}(10r_i)^s \\
&\leq \sum_{i=1}^{m} \alpha(s) r^s + \sum_{i=m+1}^{+\infty} 5^s \alpha(s) r^s \\
&\leq \frac{1}{t} \sum_{i=1}^{m} \mathcal{H}^s(B(x_i, r_i) \cap E) + \frac{5^s}{t} \sum_{i=m+1}^{+\infty} \mathcal{H}^s(B(x_i, r_i) \cap E) \\
&\leq \frac{1}{t} \mathcal{H}^s(U \cap E) + \frac{5^s}{t} \mathcal{H}^s\left(\bigcup_{i=m+1}^{+\infty} B(x_i, r_i) \cap E \right).
\end{aligned}$$

Note that this holds for each $m = 1, 2, \ldots$. Thus taking the limit as $m \to \infty$ gives

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

Letting $\delta \to 0$, we obtain

$$\mathcal{H}^s(B_t) \leq \frac{1}{t}(\mathcal{H}^s(B_t) + \epsilon),$$

and then taking the limit as $\epsilon \to 0$ gives

$$\mathcal{H}^s(B_t) \leq \frac{1}{t} \mathcal{H}^s(B_t).$$

Since $\mathcal{H}^s(B_t) \leq \mathcal{H}^s(E) < +\infty$, this implies that $\mathcal{H}^s(B_t) = 0$ for each $t > 1$, as required.

(ii) We now show that

$$\limsup_{r \to 0} \frac{\mathcal{H}^s_\infty(B(x,r) \cap E)}{\alpha(s) r^s} \geq \frac{1}{2^s}$$

for $\mathcal{H}^s$−a.e. $x \in E$.

For any $\delta > 0$ and $0 < \tau < 1$, denote by $E(\delta, \tau)$ the set of all points $x \in E$ such that

$$\mathcal{H}^s_\delta(C \cap E) \le \frac{\alpha(s)}{2^s} \tau (\operatorname{diam} C)^s,$$

whenever $C \subseteq \mathbb{R}^n$, $x \in C$, and $\operatorname{diam} C \le \delta$. Then if $\{C_i\}_{i=1}^{+\infty} \subseteq \mathbb{R}^n$ with $\operatorname{diam} C_i \le \delta$, $E(\delta,\tau) \subseteq \cup_{i=1}^{+\infty} c_i$, and $C_i \cap E(\delta,\tau) \neq \emptyset$, we have

$$\mathcal{H}^s_\delta(E(\delta,\tau)) \le \sum_{i=1}^{+\infty} \mathcal{H}^s_\delta(C_i \cap E(\delta,\tau)) \le \tau \sum_{i=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} C_i)^s.$$

Taking the infimum over all such covers $\{C_i\}_{i=1}^{+\infty}$ of $E(\delta,\tau)$, we see that

$$\mathcal{H}^s_\delta(E(\delta,\tau)) \le \tau \mathcal{H}^s_\delta(E(\delta,\tau)),$$

and so $\mathcal{H}^s_\delta(E(\delta,\tau)) = 0$, since $0 < \tau < 1$ and $\mathcal{H}^s_\delta(E(\delta,\tau)) \le \mathcal{H}^s_\delta(E) \le \mathcal{H}^s(E) < +\infty$. In particular,

$$\mathcal{H}^s(E(1-\delta,\delta)) = 0 \tag{2.3.1}$$

{eq:2.3-1

for any $0 < \delta < 1$. Now if $x \in E$ and

$$\limsup_{r \to 0} \frac{\mathcal{H}^s_\infty(B(x,r) \cap E)}{\alpha(s) r^s} < \frac{1}{2^s},$$

there exists $\delta > 0$ such that

$$\frac{\mathcal{H}^s_\infty(B(x,r) \cap E)}{\alpha(s) r^s} < \frac{1-\delta}{2^s} \tag{2.3.2}$$

{eq:2.3-2

for all $0 < r \le \delta$. Thus if $x \in C$ and $\operatorname{diam} C \le \delta$,

$$\begin{aligned}\mathcal{H}^s_\delta(C \cap E) &= \mathcal{H}^s_\infty(C \cap E) \\ &\le \mathcal{H}^s_\infty(B(x, \operatorname{diam} C) \cap E) \\ &\le (1-\delta)\frac{\alpha(s)}{2^s}(\operatorname{diam} C)^s,\end{aligned}$$

by (2.3.2). Consequently $x \in E(\delta, 1-\delta)$, and it follows

$$\left\{ x \in E : \limsup_{r \to 0} \frac{\mathcal{H}^s_\infty(B(x,r) \cap E)}{\alpha(s) r^s} < \frac{1}{2^s} \right\} \subseteq \left\{ \bigcup_{k=2}^{+\infty} E\left(\frac{1}{k}, 1 - \frac{1}{k}\right) \right\}.$$

But since the RHS has $\mathcal{H}^s-$measure zero by (2.3.1), this proves (ii).

(iii) Since $\mathcal{H}^s(B(x,r) \cap E) \ge \mathcal{H}^s_\infty(B(x,r) \cap E)$ for any $x \in E$ and $r > 0$, (ii) immediately gives the required lower estimate

$$\limsup_{r \to 0} \frac{\mathcal{H}^s(B(x,r) \cap E)}{\alpha(s) r^s} \ge \frac{1}{2^s}.$$

The proof is complete. □

2.4. **Hausdorff Measure and Elementary Properties of Functions.** We establish some properties relating the behavior of certain functions and Hausdorff measure.

. . . .

2.4.1. *Hausdorff Measure and Lipschitz Mappings.*

Definition 2.4.1 (Lipschitz)**.** *A function $F : \mathbb{R}^n \to \mathbb{R}^m$ is called Lipschitz if there exists a constant $C > 0$ such that*

$$|f(x) - f(y)| \le C|x - y|$$

for all $x, y \in \mathbb{R}^n$.

Definition 2.4.2 (Lipschitz Constant)**.** *We define the Lipschitz constant of a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ by*

$$\operatorname{Lip}(f) := \sup_{\substack{x,y\in\mathbb{R}^n\\ x\neq y}} \frac{|f(x) - f(y)|}{|x - y|}.$$

Note that for any Lipschitz function f,

$$|f(x) - f(y)| \le \operatorname{Lip}(f)|x - y|.$$

t2.4-1 **Theorem 2.4.1.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $A \subseteq \mathbb{R}^n$, $0 \le s < +\infty$. Then*

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

Proof. Fix $\delta > 0$ and choose sets $\{C_i\}_{i=1}^{+\infty} \subseteq \mathbb{R}^n$ such that $\operatorname{diam} C_i \le \delta$, $A \subseteq \cup_{i=1}^{+\infty} C_i$. Then

$$\operatorname{diam} f(C_i) \le \operatorname{Lip}(f) \operatorname{diam} C_i \le \delta \operatorname{Lip}(f),$$

and $f(A) \subseteq f(\cup_{i=1}^{+\infty} C_i) = \cup_{i=1}^{+\infty} f(C_i)$. Thus

$$\begin{aligned}\mathcal{H}^s_{\delta \operatorname{Lip}(f)}(f(A)) &\le \sum_{i=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} f(C_i))^s \\ &\le (\operatorname{Lip}(f))^s \sum_{i=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} C_i)^s.\end{aligned}$$

Taking the infimum over all such sets $\{C_i\}_{i=1}^{+\infty}$ which cover A, we find on the RHS

$$\mathcal{H}^s_{\delta \operatorname{Lip}(f)}(f(A)) \le (\operatorname{Lip}(f))^s \mathcal{H}^s_\delta(A).$$

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

$$\mathcal{H}^s(f(A)) \le (\operatorname{Lip}(f))^s \mathcal{H}^s(A),$$

as required. The proof is complete. □

c2.4-1 **Corollary 2.4.1.** *Suppose that $n > k$. Let $P : \mathbb{R}^n \to \mathbb{R}^k$ be the usual projection, $A \subseteq \mathbb{R}^n$, $0 \le s < +\infty$. Then*

$$\mathcal{H}^s(P(A)) \le \mathcal{H}^s(A).$$

Proof. Since P is the standard projection map from $\mathbb{R}^n$ to $\mathbb{R}^k$, $\operatorname{Lip}(P) = 1$. Applying the above theorem (cf. (2.4.1)) gives the required estimate. □

. . . .

2.4.2. *Graphs of Lipschitz Functions.*

Definition 2.4.3 (Graph). *For $f : \mathbb{R}^n \to \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, we define the graph $\Gamma(f; A)$ of f over A by*

$$\Gamma(f; A) := \{(x, f(x)) : x \in A\} \subseteq \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}.$$

t2.4-2 **Theorem 2.4.2.** *Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$, $\mathcal{L}^n(A) > 0$.*

(i) Then $\mathcal{H}_{\dim}(\Gamma(f; A)) \geq n$;

(ii) If f is Lipschitz, then $\mathcal{H}_{\dim}(\Gamma(f; A)) = n$.

Remark. *We thus see that the graph of a Lipschitz function f has the expected Hausdorff dimension (think of a continuous function $f : \mathbb{R} \to \mathbb{R}$). We will see from the Area Formula that $\mathcal{H}^s(\Gamma(f; A))$ can be computed according to the usual rules of calculus.*

Proof.

(i). Let $P : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be the usual projection. Then by (2.4.1) c2.4-1,

$$\mathcal{H}^n(\Gamma(f; A)) \geq \mathcal{H}^n(A) > 0.$$

Thus $\mathcal{H}^n(\Gamma(f; A)) > 0$, so that $\mathcal{H}_{\dim}(\Gamma(f; A)) \geq n$.

(ii). Let Q denote any cube in $\mathbb{R}^n$ of side length 1. Subdivide Q into k^n subcubes $\{Q_1, \dots, Q_{k^n}\}$ of side length $\frac{1}{k}$. Note that $\operatorname{diam} Q_i = \frac{\sqrt{n}}{k}$ for each $i = 1, \dots, k^n$. Define

$$a_j^i := \min_{x \in Q_j} f^i(x), \quad b_j^i := \max_{x \in Q_j} f^i(x),$$

where $i = 1, \dots, m$ and $j = 1, \dots, k^n$. Since f is Lipschitz,

$$|b_j^i - a_j^i| \leq \operatorname{Lip}(f) \operatorname{diam} Q_j = \operatorname{Lip}(f) \frac{\sqrt{n}}{k}.$$

For each $j = 1, \dots, k^n$, put

$$C_j := Q_j \times \prod_{i=1}^{m} (a_j^i, b_j^i).$$

Then

$$\Gamma(f; Q_j \cap A) = \{(x, f(x)) : x \in Q_j \cap A\} \subseteq C_j,$$

and $\operatorname{diam} C_j \leq \frac{C}{k}$ for some constant $C > 0$. Since

$$\Gamma(f; A \cap Q) = \Gamma(f; A \cap \cup_{j=1}^{k_n} Q_j) = \bigcup_{j=1}^{k_n} \Gamma(f; A \cap Q_j) \subseteq \bigcup_{j=1}^{j_n} C_j,$$

we have by monotonicity

$$\begin{aligned} \mathcal{H}^n_{C/k}(G(f; A \cap Q)) &\leq \sum_{j=1}^{k_n} \frac{\alpha(n)}{2^n} (\operatorname{diam} C_j)^n \\ &\leq \frac{k^n \alpha(n)}{2^n} \left(\frac{C}{k} \right)^n = \frac{C^n \alpha(n)}{2^n}. \end{aligned}$$

Then upon letting $k \to +\infty$, we find $\mathcal{H}^n(\Gamma(f; A \cap Q)) < +\infty$, and so $\mathcal{H}_{\dim}(\Gamma(f; A \cap Q)) \leq n$. Recall that this estimate is valid for each cube $Q \subseteq \mathbb{R}^n$ of side length 1. Consequently $\mathcal{H}_{\dim}(\Gamma(f; A)) \leq n$. Applying (i), it follows $\mathcal{H}_{\dim}(\Gamma(f; A)) = n$. The proof is complete. $\square$

. . . .

2.4.3. *The Set Where an Integrable Function is Large.* If a function f is locally integrable, we can estimate the Hausdorff measure of the set where f is locally large.

t2.4-3 **Theorem 2.4.3.** *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$, let $0 \le s < n$, and define*

$$\Lambda_s := \left\{ x \in \mathbb{R}^n : \limsup_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) > 0. \right\}$$

Then

$$\mathcal{H}^s(\Lambda_s) = 0.$$

Proof. We may as well assume that $f \in L^1(\mathbb{R}^n)$. By the Lebesgue Differentiation Theorem,

$$\lim_{r \to 0} \fint_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) = |f(x)|$$

for $\mathcal{L}^n-$a.e. $x \in \mathbb{R}^n$, and thus

$$\lim_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) = \lim_{r \to 0} \alpha(n) r^{n-s} \fint_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) = \lim_{r \to 0} \alpha(n) r^{n-s} |f(x)| = 0$$

for $\mathcal{L}^n-$a.e. $x \in \mathbb{R}^n$, since $0 \le s < n$. Hence

$$\mathcal{L}^n(\Lambda_s) = 0.$$

Fix $\epsilon > 0$, $\delta > 0$, $\sigma > 0$. Since f is $\mathcal{L}^n-$integrable, there exists $\eta > 0$ such that $\mathcal{L}^n(\Omega) \le \eta$ implies

$$\int_\Omega |f(x)| \, d\mathcal{L}^n(x) < \sigma.$$

Define

$$\Lambda_s^\epsilon := \left\{ x \in \mathbb{R}^n : \limsup_{r \to 0} \frac{1}{r^s} \int_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) > \epsilon \right\}.$$

By the above analysis,

$$\mathcal{L}^n(\Lambda_s^\epsilon) = 0.$$

Thus there exists an open set $\Omega \subseteq \mathbb{R}^n$ such that $\Lambda_s^\epsilon \subseteq \Omega$ and $\mathcal{L}^n(\Omega) < \eta$. Put

$$\mathcal{F} := \left\{ B(x,r) : x \in \Lambda_s^\epsilon, 0 < r < \delta, B(x,r) \subseteq \Omega, \int_{B(x,r)} |f(y)| \, d\mathcal{L}^n(y) > \epsilon r^s \right\}.$$

By the Vitali Covering Lemma, there exists a countable family $\{B(x_i, r_i)\}_{i=1}^{+\infty}$ of disjoint balls in $\mathcal{F}$ such that

$$\Lambda_s^\epsilon \subseteq \bigcup_{i=1}^{+\infty} B(x_i, 5r_i).$$

We thus compute

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(\Lambda_s^\epsilon) &\le \sum_{i=1}^{+\infty} \frac{\alpha(s)}{2^s} (\operatorname{diam} B(x_i, 5r_i))^s \le \sum_{i=1}^{+\infty} \alpha(s)(5r_i)^s \\
&\le \frac{\alpha(s) 5^s}{\epsilon} \sum_{i=1}^{+\infty} \int_{B(x_i,r_i)} |f(y)| \, d\mathcal{L}^n(y) \\
&\le \frac{\alpha(s) 5^s}{\epsilon} \int_\Omega |f(y)| \, d\mathcal{L}^n(y)
\end{aligned}$$

$$\leq \frac{\alpha(s)5^s}{\epsilon}\sigma.$$

Taking the limit as $\delta \to 0$, we have

$$\mathcal{H}^s(\Lambda_s^\epsilon) \leq \frac{\alpha(s)5^s}{\epsilon}\sigma,$$

and then upon sending $\sigma \to 0$ we obtain

$$\mathcal{H}^s(\Lambda_s^\epsilon) = 0.$$

Since $\epsilon > 0$ was arbitrary, it follows

$$\mathcal{H}^s(\Lambda_s) = 0.$$

The proof is complete. □

3. Area and Coarea Formulas

3.1. Lipschitz Functions, Rademacher's Theorem.

Definition 3.1.1 (Lipschitz)**.** *Let $A \subseteq \mathbb{R}^n$. A function $f : A \to \mathbb{R}^m$ is called Lipschitz provided that*

$$|f(x) - f(y)| \leq C|x - y| \tag{3.1.1}$$

{eq:3.1-1

for some constant $C > 0$ and all $x, y \in A$. The smallest constant C such that (3.1.1) *holds for all $x, y \in A$ is denoted*

$$\mathrm{Lip}(f) := \sup\left\{\frac{|f(x) - f(y)|}{|x - y|} : x, y \in A, x \neq y\right\}.$$

Definition 3.1.2 (Locally Lipschitz)**.** *A function $f : A \to \mathbb{R}^m$ is called locally Lipschitz if for each compact set $K \subseteq A$, there exists a constant $C_K > 0$ such that*

$$|f(x) - f(y)| \leq C_K|x - y|$$

for all $x, y \in K$.

t3.1-1 **Theorem 3.1.1** (Extension of Lipschitz Functions)**.** *Assume that $A \subseteq \mathbb{R}^n$, and let $f : A \to \mathbb{R}^m$ be Lipschitz. There exists a Lipschitz function $\overline{f} : \mathbb{R}^n \to \mathbb{R}^m$ such that*

- *(i)* $\overline{f} = f$ *on* A;
- *(ii)* $\mathrm{Lip}(\overline{f}) \leq \sqrt{m}\,\mathrm{Lip}(f)$.

Proof.

(i). First assume that $f : A \to \mathbb{R}$. Define

$$\overline{f}(x) := \inf_{x \in A}\{f(a) + \mathrm{Lip}(f)|x - a|\}.$$

If $b \in A$, then we have $\overline{f}(b) = f(b)$. This follows because if $b \in A$, then

$$\overline{f}(b) \leq f(b) + \mathrm{Lip}(f)|b - b| = f(b).$$

On the other hand, for all $a \in A$, we've:

$$f(a) + \mathrm{Lip}(f)|b - a| \geq f(a) + \frac{f(b) - f(a)}{|b - a|}|b - a| = f(b).$$

Taking the infimum over all $a \in A$ on the LHS thus gives $\overline{f}(b) \geq f(b)$. Now if $x, y \in \mathbb{R}^n$, then

$$\begin{aligned}\overline{f}(x) &\leq \inf_{a \in A}\{f(a) + \mathrm{Lip}(f)(|x - y| + |y - a|)\}\\ &= \inf_{a \in A}\{f(a) + \mathrm{Lip}(f)|y - a|\} + \mathrm{Lip}(f)|x - y|\\ &= \overline{f}(y) + \mathrm{Lip}(f)|x - y|.\end{aligned}$$

Similarly

$$\overline{f}(y) \leq \overline{f}(x) + \mathrm{Lip}(f)|x - y|.$$

Therefore

$$\frac{|\overline{f}(x) - \overline{f}(y)|}{|x - y|} \leq \mathrm{Lip}(f)$$

for all $x, y \in A$. This proves the result for functions $f : A \to \mathbb{R}$.

(ii). In the general case $f : A \to \mathbb{R}^m$, $f = (f^1, \ldots, f^m)$, define $\overline{f} := (\overline{f}^1, \ldots, \overline{f}^m)$, where $\overline{f}^i$, $i = 1, \ldots, m$, are defined as in (i). Then

$$|\overline{f}(x) - \overline{f}(y)|^2 = \sum_{i=1}^{m} \left|\overline{f}^i(x) - \overline{f}^i(y)\right|^2 \leq m(\mathrm{Lip}(f))^2|x-y|^2.$$

Taking square roots,

$$\overline{f}(x) - \overline{f}(y) \leq \sqrt{m}\,\mathrm{Lip}(f)|x-y|,$$

as required. The proof is complete. □

Remark. *In fact there exists an extension $\overline{f}$ of f with $\mathrm{Lip}(\overline{f}) = \mathrm{Lip}(f)$. This is Kirszbraun's Theorem.*

We now prove Rademacher's Theorem, which states that a locally Lipschitz function is differentiable $\mathcal{L}^n-$a.e. Note that the inequality

$$|f(x) - f(y)| \leq \mathrm{Lip}(f)|x-y|$$

says nothing about the possibility of locally approximating f by a linear map.

Definition 3.1.3 (Differentiable)**.** *The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be differentiable at $x \in \mathbb{R}^n$ if there exists a linear mapping*

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

such that

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

or, equivalently,

$$f(y) = f(x) + L(x-y) + o(|y-x|), \quad y \to x.$$

Remark.

(i) Note that this is actually the definition of the Fréchet derivative.

(ii) If such a linear mapping L exists, it is unique, and we write

$$Df(x)$$

for L. We call $Df(x)$ the derivative of f at x.

t3.1-2 **Theorem 3.1.2** (Rademacher's Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. Then f is differentiable $\mathcal{L}^n-$a.e.*

Proof.

(i). We may assume that $m = 1$, for otherwise, repeat the below argument m times. Since differentiability is a local property, we may as well also suppose that f is Lipschitz.

(ii). Fix any $v \in \mathbb{R}^n$ with $|v| = 1$, and for any $x \in \mathbb{R}^n$, define the Gateaux derivative

$$D_v f(x) := \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

at x, provided that this limit exists.

(iii). We show that $D_v f(x)$ exists for $\mathcal{L}^n-$a.e. $x \in \mathbb{R}^n$. Since f is continuous,

$$\overline{D}_v f(x) = \limsup_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

$$= \lim_{k\to+\infty} \sup_{\substack{0<|t|<\frac{1}{k} \\ t\in\mathbb{Q}}} \frac{f(x+tv)-f(x)}{t}$$

is Borel measurable, as is

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

Thus

$$\begin{aligned} A_v &:= \{x\in\mathbb{R}^n : D_v f(x) \text{ does not exist}\} \\ &= \left\{x\in\mathbb{R}^n : \underline{D}_v f(x) < \overline{D}_v f(x)\right\}, \end{aligned}$$

being the complement of the set of all points of which the pointwise limit of measurable functions exists, is Borel measurable.

Now, for each $x, v \in \mathbb{R}^n$ with $|v| = 1$, define $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(t) := f(x+tv).$$

Note that for any $t \in \mathbb{R}$,

$$\begin{aligned} |\phi(t)-\phi(s)| &= |f(x+tv)-f(x+sv)| \le \operatorname{Lip}(f)|(x+tv)-(x+sv)| \\ &= \operatorname{Lip}(f)|t-s|, \end{aligned}$$

so that ϕ is Lipschitz. Therefore ϕ is absolutely continuous, and thus differentiable $\mathcal{L}^1-$a.e. Thus for any line L parallel to v, the set of all points on L such that f is not differentiable has Lebesgue measure zero. That is,

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

for each line L parallel to v. Thus the Fubini–Tonelli Theorem implies

$$\mathcal{L}^n(A_v) = 0,$$

as required.

(iv). Noting that

$$\frac{\partial}{\partial x_j} f(x) = D_{e_j} f(x) = \lim_{t\to 0} \frac{f(x+te_j)-f(x)}{t}$$

for each $j = 1, \ldots, n$, we have by (iii) that

$$\nabla f(x) = \left(\frac{\partial}{\partial x_1} f(x), \ldots, \frac{\partial}{\partial x_n} f(x)\right)$$

exists for $\mathcal{L}^n-$a.e. $x \in \mathbb{R}^n$.

(v). Next we show that $D_v f(x) = v \cdot \nabla f(x)$ for $\mathcal{L}^n-$a.e. $x \in \mathbb{R}^n$. Let $\zeta \in \mathcal{C}_c^\infty(\mathbb{R}^n)$. Then

$$\begin{aligned} \int_{\mathbb{R}^n} \left[\frac{f(x+tv)-f(x)}{t}\right] \zeta(x)\, dx &= \frac{1}{t}\left[\int_{\mathbb{R}^n} f(x+tv)\zeta(x)\, dx - \int_{\mathbb{R}^n} f(x)\zeta(x)\, dx\right] \\ &= \frac{1}{t}\left[\int_{\mathbb{R}^n} f(x)\zeta(x-tv)\, dx - \int_{\mathbb{R}^n} f(x)\zeta(x)\, dx\right] \\ &= -\int_{\mathbb{R}^n} f(x)\left[\frac{\zeta(x)-\zeta(x-tv)}{t}\right]\, dx. \end{aligned}$$

This is the integration by parts formula for difference quotients. Let $t = \frac{1}{k}$ for $k = 1, 2, \ldots,$ in the above equality and note that

$$\frac{|f(x + \frac{1}{k}v) - f(x)|}{\frac{1}{k}} \leq \operatorname{Lip}(f).$$

Thus, by Lebesgue's Dominated Convergence Theorem, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\zeta(x)\, dx &\overset{LDC}{=} -\int_{\mathbb{R}^n} f(x) D_v\zeta(x)\, dx \\
&= -\sum_{j=1}^{n} v_i \int_{\mathbb{R}^n} f(x) \frac{\partial}{\partial x_j}\zeta(x)\, dx \\
&= \sum_{j=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial}{\partial x_j} f(x)\zeta(x)\, dx \\
&= \int_{\mathbb{R}^n} (v \cdot \nabla f(x))\zeta(x)\, dx,
\end{aligned}$$

where we have used integration by parts and the partial derivatives on f are understood in the a.e. sense. Since the above equality holds for every $\zeta \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, we have $D_v f = v \cdot \nabla f$ $\mathcal{L}^n$–a.e.

(vi). Choose $\{v_k\}_{k=1}^{+\infty}$ to be a countable, dense subset of $\partial B(0,1)$. Set

$$A_k := \{x \in \mathbb{R}^n : D_{v_k} f(x),\ \nabla f(x) \text{ exist and } D_{v_k} f(x) = v_k \cdot \nabla f(x)\}$$

for each $k \in \mathbb{N}$. Note that by (iii)-(v), $\mathcal{L}^n(\mathbb{R}^n \setminus A_k) = 0$ for each $k \in \mathbb{N}$. Define

$$A := \bigcap_{k=1}^{+\infty} A_k$$

and observe that

$$\mathcal{L}^n(\mathbb{R}^n \setminus A) = \mathcal{L}^n(\mathbb{R}^n \setminus \cap_{k=1}^{+\infty} A_k) = \mathcal{L}^n(\cup_{k=1}^{+\infty}(\mathbb{R}^n \setminus A_k)) = 0.$$

(vii). We now show that f is differentiable at each point $x \in A$. Fix any $x \in A$. Choose $v \in \partial B(0,1)$, $t \in \mathbb{R}$, $t \neq 0$, and write

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

Then if $w \in \partial B(0,1)$, we have

$$\begin{aligned}
|Q(x,v,t) - Q(x,w,t)| &= \left| \frac{f(x+tv) - f(x+tw)}{t} - (v-w)\cdot \nabla f(x) \right| \\
&\leq \left| \frac{f(x+tv) - f(x+tw)}{t} \right| + |(v-w)\cdot \nabla f(x)| \\
&\leq \operatorname{Lip}(f)|v-w| + |\nabla f(x)||v-w| \\
&\leq (1+\sqrt{n}) \operatorname{Lip}(f)|v-w|. \qquad (3.1.2)
\end{aligned}$$

{eq:3.1-2

Fix $\epsilon > 0$ and choose $N \in \mathbb{N}$ so large that if $v \in \partial B(0,1)$, then

$$|v - v_k| \leq \frac{\epsilon}{2(1+\sqrt{n}) \operatorname{Lip}(f)}$$

for some $k = 1, \dots, N$. Note that since $x \in A$,

$$\begin{aligned} \lim_{t\to 0} Q(x, v_k, t) &= \lim_{t\to 0} \left\{ \frac{f(x+tv_k) - f(x)}{t} - v_k \cdot \nabla f(x) \right\} \\ &= D_{v_k} f(x) - v_k \cdot \nabla f(x) \\ &= 0 \end{aligned}$$

for each $k = 1, \dots, N$. Thus there exists $\delta > 0$ so that for all $0 < |t| < \delta$,

$$|Q(x, v_k, t)| < \frac{\epsilon}{2} \tag{3.1.3}$$ {eq:3.1-3

holds for each $k = 1, \dots, N$. Consequently for each $v \in \partial B(0,1)$ there exists $k \in \{1, \dots, k\}$ such that

$$\begin{aligned} |Q(x,v,t)| &\le |Q(x,v,t) - Q(x,v_k,t)| + |Q(x,v_k,t)| \\ &< (1+\sqrt{n}) \operatorname{Lip}(f)|v - v_k| + \frac{\epsilon}{2} \\ &< \epsilon, \end{aligned}$$

by (3.1.2) and (3.1.3), provided that $0 < |t| < \delta$. Note that this is the same $\delta > 0$ for all $v \in \partial B(0,1)$.

Now choose any $x, y \in \mathbb{R}^n$, $y \neq x$. Write

$$v := \frac{y-x}{|y-x|},$$

so that $y = x + tv$, where $t := |x - y|$. Then

$$\begin{aligned} |f(y) - f(x) - \nabla f(x) \cdot (y-x)|| &= |f(x+tv) - f(x) - \nabla f(x) \cdot tv| \\ &= |Q(x,t,v)||t| \\ &< \epsilon |t|, \end{aligned}$$

so that

$$f(y) - f(x) - \nabla f(x) \cdot (y-x) = o(t) = o(|x-y|), \quad y \to x.$$

Hence, f is differentiable at x, with

$$Df(x) = \nabla f(x).$$

The proof is complete. □

c3.1-1 **Corollary 3.1.1.**

(i) Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz, and

$$\mathcal{Z} := \{x \in \mathbb{R}^n : f(x) = 0\}.$$

Then $Df(x) = 0$ for $\mathcal{L}^n$–a.e. $x \in \mathcal{Z}$.

(ii) Let $f, g := \mathbb{R}^n \to \mathbb{R}^n$ be locally Lipschitz, and

$$Y := \{x \in \mathbb{R}^n : g(f(x)) = x\}.$$

Then

$$Dg(f(x))Df(x) = I$$

for $\mathcal{L}^n$–a.e. $x \in Y$.

Proof.

(i). We may assume that $m = 1$ in (i), otherwise, repeat the following argument m times.

(ii). Choose $x \in \mathcal{Z}$ so that $Df(x)$ exists, and

$$\lim_{r \to 0} \frac{\mathcal{L}^n(\mathcal{Z} \cap B(x,r))}{\mathcal{L}^n(B(x,r))} = 1. \tag{3.1.4}$$

{eq:3.1-4

Note that this holds for $\mathcal{L}^n$−a.e. $x \in \mathcal{Z}$. Since $x \in \mathcal{Z}$, it follows

$$f(y) = Df(x) \cdot (y - x) + o(|y - x|). \tag{3.1.5}$$

{eq:3.1-5

By contradiction, suppose that $Df(x) = \alpha \neq 0$, and set

$$S := \left\{ v \in \partial B(0,1) : \alpha \cdot v \geq \frac{1}{2}|\alpha| \right\}.$$

Note that S is nonempty, for otherwise $Df(x) = 0$. Now for each $v \in S$ and $t > 0$, set $y := x + tv$ in (3.1.5) to obtain

$$\begin{aligned} f(x + tv) &= \alpha \cdot tv + o(|tv|) \\ &\geq \frac{|\alpha|}{2} t + o(t). \end{aligned}$$

Hence, there exists $\delta > 0$ such that for all $0 < t < \delta$ and all $v \in S$,

$$f(x + tv) > 0.$$

But this contradicts (3.1.4), since for all $0 < r < \delta$, $B(x,r) \cap \mathcal{Z} = \{x\}$. This proves (i).

(iii). We now show (ii). Define

$$\operatorname{dom} Df := \{x \in \mathbb{R}^n : Df(x) \text{ exists}\}$$

and

$$\operatorname{dom} Dg := \{x \in \mathbb{R}^n : Dg(x) \text{ exists}\}.$$

Put

$$X := Y \cap \operatorname{dom} Df \cap f^{-1}(\operatorname{dom} Dg).$$

Then

$$\begin{aligned} Y \setminus X &= Y \cap \left(Y^C \cup (\operatorname{dom} Df)^C \cup (f^{-1}(\operatorname{dom} Dg))^C\right) \\ &= (Y \setminus \operatorname{dom} Df) \cup (Y \setminus f^{-1}(\operatorname{dom} Dg)) \\ &\subseteq (\mathbb{R}^n \setminus \operatorname{dom} Df) \cup g(\mathbb{R}^n \setminus \operatorname{dom} Dg). \end{aligned} \tag{3.1.6}$$

{eq:3.1-6

This follows since if $x \in Y \setminus f^{-1}(\operatorname{dom} Dg)$, then $f(x) \in f(Y) \subseteq \mathbb{R}^n$, and $f(x) \notin \operatorname{dom} Dg$, so that

$$f(x) \in \mathbb{R}^n \setminus \operatorname{dom} Dg.$$

Thus

$$x = g(f(x)) \in g(\mathbb{R}^n \setminus \operatorname{dom} Dg.)$$

By Rademacher's Theorem (cf. (3.1.2)),

$$\mathcal{L}^n(\mathbb{R}^n \setminus \operatorname{dom} Df) = 0$$

and

$$\mathcal{L}^n(\mathbb{R}^n \setminus \operatorname{dom} Dg) = 0.$$

Moreover, since g is Lipschitz (cf. (2.4.1)), we have

$$\mathcal{L}^n(g(\mathbb{R}^n \setminus \operatorname{dom} Dg)) \leq (\operatorname{Lip}(g))^n \mathcal{L}^n(\mathbb{R}^n \setminus \operatorname{dom} Dg) = 0.$$

Thus, by (3.1.6),

$$\mathcal{L}^n(Y \setminus X) = 0.$$

Now if $x \in X$, $Dg(f(x))$ and $Df(x)$ exist, and so the chain rule implies

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

exists. Finally, since $(g \circ f)(x) - x = g(f(x)) - x = 0$ on Y, assertion (i) gives

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

$\mathcal{L}^n-$a.e. on Y. The proof is complete. □

3.2. **Linear Maps and Jacobians.** We first review some basic linear algebra. Our goal in this section is to define the Jacobian of a map $f : \mathbb{R}^n \to \mathbb{R}^m$.

. . . .

3.2.1. *Linear Maps.*

Definition 3.2.1 (Orthogonal Linear Map)**.** *A linear map $O : \mathbb{R}^n \to \mathbb{R}^m$ is orthogonal if*

$$Ox \cdot Oy = x \cdot y$$

for all $x, y \in \mathbb{R}^n$.

Definition 3.2.2 (Symmetric Linear Map)**.** *A linear map $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric if*

$$x \cdot Sy = Sx \cdot y$$

for all $x, y \in \mathbb{R}^n$.

Definition 3.2.3 (Diagonal Linear Map)**.** *A linear map $D : \mathbb{R}^n \to \mathbb{R}^n$ is diagonal if there exist $d_1, \ldots, d_n \in \mathbb{R}$ such that*

$$Dx = (d_1 x_1, \ldots, d_n x_n)$$

for all $x \in \mathbb{R}^n$.

Definition 3.2.4 (Adjoint)**.** *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. The adjoint of A is the linear map $A^* : \mathbb{R}^m \to \mathbb{R}^n$ defined by*

$$x \cdot A^* y = Ax \cdot y$$

for all $x \in \mathbb{R}^n, y \in \mathbb{R}^m$.

Recall that the existence of adjoints in Euclidean space with the Euclidean metric is guaranteed, and, since $\mathbb{R}^n$ is a Hilbert space under the Euclidean metric, the adjoint operator has the above form by the Riesz Representation Theorem.

t3.2-1 **Theorem 3.2.1.**

(i) $A^{**} = A$;
(ii) $(A \circ B)^* = B^* \circ A^*$;
(iii) If $O : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal, then $O^* = O^{-1}$;
(iv) If $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric, then $S^* = S$;

(v) If $S : \mathbb{R}^n \to \mathbb{R}^n$ is symmetric, there exists an orthogonal map $O : \mathbb{R}^n \to \mathbb{R}^n$ and a diagonal map $D : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S = O \circ D \circ O^{-1};$$

(vi) If $O : \mathbb{R}^n \to \mathbb{R}^m$ is orthogonal, then $n \le m$ and

$$O^* \circ O = I \quad \text{on } \mathbb{R}^n,$$
$$O \circ O^* = I \quad \text{on } O(\mathbb{R}^n).$$

Proof.
(i). Since the dot product is symmetric, we have for all $x, y \in \mathbb{R}^n$ that

$$\begin{aligned} x \cdot (A^{**}y) &= x \cdot (A^*)^* y = A^*x \cdot y = y \cdot A^*x = Ay \cdot x \\ &= x \cdot Ay. \end{aligned}$$

Since this is for all $x \in \mathbb{R}^n$, assertion (i) follows.
(ii). For any $x, y \in \mathbb{R}^n$,

$$\begin{aligned} x \cdot (A \circ B)^* y &= (A \circ B)x \cdot y = A(Bx) \cdot y = Bx \cdot A^*y \\ &= x \cdot B^*(A^*y). \end{aligned}$$

This is for all $x \in \mathbb{R}^n$, so this proves (ii).
(iii). Let $x, y \in \mathbb{R}^n$. Then

$$x \cdot y = Ox \cdot Oy = x \cdot O^*(Oy),$$

and

$$x \cdot y = O(O^{-1}x) \cdot y = O^{-1}x \cdot O^*y = x \cdot O(O^*y).$$

This shows $O^* = O^{-1}$.
(iv). If $x, y \in \mathbb{R}^n$, then

$$x \cdot Sy = Sx \cdot y = x \cdot S^*y,$$

and since this is for all $x \in \mathbb{R}^n$, assertion (iv) follows. □

t3.2-2 **Theorem 3.2.2** (Polar Decomposition). *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping.*

(i) If $n \le m$, there exists a symmetric map $S : \mathbb{R}^n \to \mathbb{R}^n$ and an orthogonal map $O : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$L = O \circ S.$$

(ii) If $n \ge m$, there exists a symmetric map $S : \mathbb{R}^m \to \mathbb{R}^m$ and an orthogonal map $O : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$L = S \circ O^*.$$

Proof.
(i). First suppose $n \le m$. Consider the mapping $C := L^* \circ L : \mathbb{R}^n \to \mathbb{R}^n$. Now for any $x, y \in \mathbb{R}^n$,

$$\begin{aligned} Cx \cdot y &= (L^* \circ L)x \cdot y = L^*(Lx) \cdot y = Lx \cdot Ly = x \cdot L^*(Ly) = x \cdot (L^* \circ L)y \\ &= x \cdot Cy, \end{aligned}$$

and also

$$Cx \cdot x = (L^* \circ L)x \cdot x = L^*(Lx) \cdot x = Lx \cdot Lx \ge 0.$$

Thus C is symmetric and positive semidefinite. Hence there exist $\mu_1, \ldots, \mu_n \geq 0$ and an orthonormal basis $\{x_k\}_{k=1}^n$ of $\mathbb{R}^n$ such that

$$Cx_k = \mu_k x_k,$$

$k = 1, \ldots, n$. Write $\mu_k := \lambda_k^2$, $\lambda_k \geq 0$, $k = 1, \ldots, n$.

(ii). We show that there exists an orthonormal set $\{z_k\}_{k=1}^n$ in $\mathbb{R}^m$ such that

$$Lx_k = \lambda_k z_k,$$

$k = 1, \ldots, n$. To see this, if $\lambda_k \neq 0$, define

$$z_k := \frac{1}{\lambda_k} L x_k.$$

Then if $\lambda_k, \lambda_l \neq 0$,

$$\begin{aligned} z_k \cdot z_l &= \frac{1}{\lambda_k} L x_k \cdot \frac{1}{\lambda_l} L x_l = \frac{1}{\lambda_k \lambda_l} L x_k \cdot L x_l = \frac{1}{\lambda_k \lambda_l} x_k \cdot L^*(L x_l) = \frac{1}{\lambda_k \lambda_l} x_k \cdot C x_l \\ &= \frac{\lambda_l^2}{\lambda_k \lambda_l} x_k \cdot x_l \\ &= \frac{\lambda_l}{\lambda_k} \delta_{kl}, \end{aligned}$$

by (i) and the fact that $\{x_k\}_{k=1}^n$ is an orthonormal set. Thus the set $\{z_k : \lambda_k \neq 0\}$ is orthonormal. If $\lambda_k = 0$, define z_k to be any unit vector such that the set $\{z_k\}_{k=1}^n$ is orthonormal, applying the Gram–Schmidt process if necessary.

(iii). Define $S : \mathbb{R}^n \to \mathbb{R}^n$ by

$$Sx_k := \lambda_k x_k,$$

$k = 1, \ldots, n$ and $O : \mathbb{R}^n \to \mathbb{R}^m$ by

$$Ox_k := z_k,$$

$k = 1, \ldots, n$. Then

$$(O \circ S)x_k = O(S_k) = O(\lambda_k)x_k = \lambda_k O x_k = \lambda_k z_k = Lx_k,$$

and, since $\{x_k\}_{k=1}^n$ is a basis for $\mathbb{R}^n$,

$$L = O \circ S.$$

Notice that the mapping S is clearly symmetric. Moreover, O is orthogonal because

$$Ox_k \cdot Ox_l = z_k \cdot z_l = \delta_{kl} = x_k \cdot x_l.$$

This proves assertion (i) of the theorem.

(iv). To prove assertion (ii), we apply assertion (i) to L^* and apply (3.2.1) to obtain

$$L^* = (O \circ S)^* = S^* \circ O^* = S \circ O^*.$$

The proof is complete. □

We now define the Jacobian of a linear map.

Definition 3.2.5 (Jacobian). *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.*

(i) If $n \leq m$, write $L = O \circ S$ (cf. (3.2.2)), and we define the Jacobian of L to be

$$[\![L]\!] := |\det S|;$$

(ii) If $n \geq m$, write $L = S \circ O^*$ (cf. (3.2.2)), and we define the Jacobian of L to be

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

Remark.

(i) It will follow from Theorem (3.2.3) below that the definition of $[\![L]\!]$ is independent of the particular choices of O and S.

(ii) Note that if, say, $n \leq m$, then $L = O \circ S$ implies

$$L^* = (O \circ S)^* = S^* \circ O^* = S \circ O^*.$$

This is the same O and S, and it clearly follows

$$[\![L]\!] = [\![L^*]\!].$$

t3.2-3 **Theorem 3.2.3.**

(i) If $n \leq m$,

$$[\![L]\!]^2 = \det(L^* \circ L);$$

(ii) If $n \geq m$,

$$[\![L]\!]^2 = \det(L \circ L^*).$$

Proof.

(i). Assume that $n \leq m$, and apply Theorem (3.2.2) to write

$$L = O \circ S$$

and

$$L^* = (O \circ S)^* = S^* \circ O^* = S \circ O^*.$$

Then

$$L^* \circ L = (S \circ O^*) \circ (O \circ S) = S \circ (O^* \circ O) \circ S = S \circ S = S^2$$

(cf. (3.2.1)). Hence,

$$\det(L^* \circ L) = \det(S^2) = (\det S)^2 = [\![L]\!],$$

as required.

(ii). The proof of (ii) is similar. The proof is complete. □

Theorem (3.2.3) provides us with a nice way to compute the Jacobian $[\![L]\!]$ of a linear map. We augment this with the Binet–Cauchy formula below.

Definition 3.2.6 ($\Lambda(m,n)$)**.** *If $n \leq m$, we define*

$$\Lambda(m,n) := \{\lambda : \{1,\ldots,n\} \to \{1,\ldots,m\} : \lambda \textit{ strictly increasing}\}.$$

Note that this is the set of all functions λ that take $\{1,\ldots,n\}$ to $\{1,\ldots,m\}$ such that $\lambda(k) > \lambda(l)$ if $k > l$, $k, l \in \{1,\ldots,n\}$.

Definition 3.2.7 (P_λ)**.** *If $n \leq m$, for each $\lambda \in \Lambda(m,n)$, we define $P_\lambda : \mathbb{R}^m \to \mathbb{R}^n$ by*

$$P_\lambda(x_1,\ldots,x_m) := (x_{\lambda(1)},\ldots,x_{\lambda(n)}).$$

We may think of P_λ as a mapping that "deletes" points from $(x_1,\ldots,x_m)$.

Remark. *For each $\lambda \in \Lambda(m,n)$, there exists an $n-$dimensional subspace*

$$S_\lambda := \operatorname{span}\{e_{\lambda(1)},\ldots,e_{\lambda(n)}\} \subseteq \mathbb{R}^m$$

such that P_λ is the projection of $\mathbb{R}^m$ onto S_λ.

t3.2-4 **Theorem 3.2.4** (Binet–Cauchy Formula). *Let $n \leq m$ and let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then*

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

Remark.

(i) To calculate $[\![L]\!]$, we compute the sums of the squares of the determinants of each $n \times n$ submatrix of the $m \times n$ matrix representing L, with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$;

(ii) This is a kind of higher dimensional version of the Pythagorean Theorem.

Proof.

(i). Identifying linear maps with their matrices with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, we write

$$L : +((l_{ij}))_{m\times n}, \quad A := L^* \circ L = ((a_{ij}))_{n\times n};$$

so that

$$a_{ij} = \sum_{k=1}^m l_{ki} l_{kj}, \quad i,j = 1,\ldots,n.$$

(ii). Note that

$$[\![L]\!]^2 = \det A = \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)},$$

where Σ denotes the set of all permutations of $\{1,\ldots,n\}$. Thus

$$\begin{aligned}
[\![L]\!]^2 &= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^n \sum_{k=1}^m l_{ki} l_{k\sigma(i)} \\
&= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\phi\in\Phi} \prod_{i=1}^n l_{\phi(i)i} l_{\phi(i)\sigma(i)},
\end{aligned}$$

where Φ denotes the set of all one–to–one mappings of $\{1,\ldots,n\}$ into $\{1,\ldots,m\}$.

(iii). Now for each $\phi \in \Phi$, we can uniquely write $\phi := \lambda\circ\theta$, where $\theta \in \Sigma$ and $\lambda \in \Lambda(m,n)$. Consequently we have

$$\begin{aligned}
[\![L]\!]^2 &= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \prod_{i=1}^n l_{\lambda\circ\theta(i),i} l_{\lambda\circ\theta(i),\sigma(i)} \\
&= \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \prod_{i=1}^n l_{\lambda(i),\theta^{-1}(i)} l_{\lambda(i),\sigma\circ\theta^{-1}(i)} \\
&= \sum_{\lambda\in\Lambda(m,n)} \sum_{\theta\in\Sigma} \sum_{\sigma\in\Sigma} \operatorname{sgn}(\sigma) \prod_{i=1}^n l_{\lambda(i),\theta(i)} l_{\lambda(i),\sigma\circ\theta(i)}.
\end{aligned}$$

Set $\rho := \sigma\circ\theta$. Then

$$[\![(]\!]L)^2 = \sum_{\lambda\in\Lambda(m,n)} \sum_{\rho\in\Sigma} \sum_{\theta\in\Sigma} \operatorname{sgn}(\theta) \operatorname{sgn}(\rho) \prod_{i=1}^n l_{\lambda(i),\theta(i)} l_{\lambda(i),\rho(i)}$$

$$= \sum_{\lambda\in\Lambda(m,n)} \left(\sum_{\theta\in\Sigma} \operatorname{sgn}(\theta) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)} \right)^2$$
$$= \sum_{\lambda\in\Lambda(m,n)} (\det(P_\lambda) \circ L)^2,$$

as required. The proof is complete. $\square$

. . . .

3.2.2. *Jacobians.* Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz mapping. By Rademacher's Theorem (cf. (3.1.2)), f is differentiable $\mathcal{L}^n$−a.e., and therefore $Df(x)$ exists and may be regarded as a linear mapping from $\mathbb{R}^n$ into $\mathbb{R}^m$ for $\mathcal{L}^n$−a.e. $x \in \mathbb{R}^n$. We recall the definition of a gradient matrix.

Definition 3.2.8 (Gradient Matrix). *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, $f = (f^1, \dots, f^m)$, we define the gradient matrix*

$$Df(x) := \begin{bmatrix} \frac{\partial}{\partial x_1} f^1(x) & \cdots & \frac{\partial}{\partial x_n} f^1(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f^m(x) & \cdots & \frac{\partial}{\partial x_n} f^m(x) \end{bmatrix}.$$

Definition 3.2.9 (Jacobian). *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, the Jacobian of f is*

$$Jf(x) := [\![Df(x)]\!], \quad \mathcal{L}^n - a.e.$$

Note that in view of Theorem (3.2.3), we have

$$(Jf(x))^2 = \det(Df(x)^* \circ Df(x)) = \det(Df(x) \circ Df(x)^*).$$

3.3. **The Area Formula.** Throughout this section we assume that

$$n \le m.$$

. . . .

3.3.1. *Preliminaries.*

l3.3-1 **Lemma 3.3.1.** *Suppose that $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, $n \le m$. Then*

$$\mathcal{H}^n(L(A)) = [\![L]\!]\mathcal{L}^n(A)$$

for all $A \subseteq \mathbb{R}^n$.

Proof.

(i). Write $L := O \circ S$, where $O : \mathbb{R}^n \to \mathbb{R}^m$ is an orthogonal map and $S : \mathbb{R}^n \to \mathbb{R}^n$ a symmetric map (cf t3.2-2 (3.2.2)). Recall that $[\![L]\!] = |\det S|$.

(ii). If $[\![L]\!] = 0$, then $\dim S(\mathbb{R}^n) \le n-1$, and so $\dim L(\mathbb{R}^n) \le n-1$. Consequently $\mathcal{H}^n(L(A)) = 0$, and the inequality is trivial.

(iii). If $[\![L]\!] > 0$, then

$$\begin{aligned}
\frac{\mathcal{H}^n(L(B(x,r)))}{\mathcal{L}^n(B(x,r))} &= \frac{\mathcal{L}^n(O^* \circ L(B(x,r)))}{\mathcal{L}^n(B(x,r))} \\
&= \frac{\mathcal{L}^n(O^* \circ O \circ S(B(x,r)))}{\mathcal{L}^n(B(x,r))} \\
&= \frac{\mathcal{L}^n(S(B(x,r)))}{\mathcal{L}^n(B(x,r))} \\
&= \frac{\mathcal{L}^n(S(B(0,1)))}{\alpha(n)} \\
&= |\det S| = [\![L]\!].
\end{aligned}$$

(iv). Define $\nu(A) := \mathcal{H}^n(L(A))$ for all $A \subseteq \mathbb{R}^n$. Then ν is a Radon measure, $\nu << \mathcal{L}^n$, and

$$D_{\mathcal{L}^n}\nu(x) = \lim_{r\to 0} \frac{\nu(B(x,r))}{\mathcal{L}^n(B(x,r))} = [\![L]\!]$$

by (iii). Thus for all Borel sets $B \subseteq \mathbb{R}^n$,

$$\mathcal{H}^n(L(B)) = [\![L]\!]\mathcal{L}^n(B).$$

Since ν and $\mathcal{L}^n$ are Radon measures, the same identity holds for all sets $A \subseteq \mathbb{R}^n$. The proof is complete. □

For the remainder of the section we assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz.

l3.3-2 **Lemma 3.3.2.** *Let $A \subseteq \mathbb{R}^n$ be $\mathcal{L}^n-$measurable. Then*

(i) $f(A)$ is $\mathcal{H}^n-$measurable;
(ii) The mapping $u \mapsto \mathcal{H}^0(A \cap f^{-1}(y))$ is $\mathcal{H}^n-$measurable on $\mathbb{R}^m$;
(iii) $\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(y))\, d\mathcal{H}^n \le (\mathrm{Lip}(f))^n \mathcal{L}^n(A)$.

Proof.

(i). We may assume without loss of generality that A is bounded.

(ii). There exist compact sets $K_i \subseteq A$ such that

$$\mathcal{L}^n(K_i) \ge \mathcal{L}^n(A) - \frac{1}{i}, \quad i = 1, \ldots, n.$$

Since $\mathcal{L}^n(A) < +\infty$ by the assumption and A is $\mathcal{L}^n-$measurable, $\mathcal{L}^n(A \setminus K_i) \le \frac{1}{i}$. Since f is continuous, $f(K_i)$ is compact and thus $\mathcal{H}^n-$measurable. Hence, $f(\cup_{i=1}^{+\infty} K_i) = \cup_{i=1}^{+\infty} f(K_i)$ is $\mathcal{H}^n-$measurable. Moreover

$$\begin{aligned}
\mathcal{H}^n\left(f(A) \setminus f\left(\bigcup_{i=1}^{+\infty} K_i\right)\right) &\le \mathcal{H}^n\left(f\left(A \setminus \bigcup_{i=1}^{+\infty} K_i\right)\right) \\
&\le (\mathrm{Lip}(f))^n \mathcal{L}^n\left(A \setminus \bigcup_{i=1}^{+\infty} K_i\right) = 0.
\end{aligned}$$

Thus $f(A)$ is $\mathcal{H}^n-$measurable. This proves (i).

(iii). Put

$$\mathcal{B}_k := \left\{ Q : Q = (a_1, b_1] \times \cdots \times (a_n, b_n], a_i := \frac{c_i}{k}, b_i := \frac{c_i+1}{k}, c_i \in \mathbb{Z}, i = 1, \ldots, n \right\},$$

and notice that

$$\mathbb{R}^n = \bigcup_{Q \in \mathcal{B}_k} Q.$$

Define

$$g_k := \sum_{Q \in \mathcal{B}_k} \mathbb{1}_{f(A \cap Q)},$$

and note that g_k is $\mathcal{H}^n-$measurable by assertion (i). Also $g_k(y)$ gives the number of cubes $Q \in \mathcal{B}_k$ such that $f^{-1}(y) \cap (A \cap Q) \neq \emptyset$. Thus

$$g_k(y) \to \mathcal{H}^0(A \cap f^{-1}(y)) \quad \text{as } k \to +\infty$$

for each $y \in \mathbb{R}^m$, and so $y \mapsto \mathcal{H}^0(A \cap f^{-1}(y))$ is $\mathcal{H}^n-$measurable.

(iv). Note that g_k as defined in (iii) satisfies

$$0 \leq g_1 \leq g_2 \leq \cdots .$$

Thus by the Monotone Convergence Theorem,

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

as required. The proof is complete. □

13.3-3 **Lemma 3.3.3.** *Let $t > 1$ and define*

$$B := \{x \in \mathbb{R}^n : Df(x) \text{ exists}, \ Jf(x) > 0\}.$$

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

(i) $B = \cup_{k=1}^{+\infty} E_k$;

(ii) $f|_{E_k}$ *is one–to–one,* $k = 1, 2, \ldots$;

(iii) For each $k = 1, 2, \ldots$, *there exists a symmetric automorphism* $T_k : \mathbb{R}^n \to \mathbb{R}^n$ *such that*

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

Proof.

(i). Fix $\epsilon > 0$ such that

$$\frac{1}{t} + \epsilon < 1 < t - \epsilon.$$

Let C be a countable dense subset of B and let S be a countable dense subset of the symmetric automorphisms of $\mathbb{R}^n$.

(ii). Then for each $c \in C$ and $T \in S$, and $i = 1, 2, \ldots$, define $E(c, T, i)$ to be the set of all $b \in B \cap B(c, \frac{1}{i})$ satisfying

$$\left(\frac{1}{t} + \epsilon\right) |Tv| \leq |Df(b)v| \leq (t - \epsilon)|Tv| \tag{3.3.1}$$

{eq3.3-1}

for all $v \in \mathbb{R}^n$ and

$$|f(a) - f(b) - Df(b) \cdot (a - b)| \leq \epsilon |T(a - b)| \tag{3.3.2}$$

{eq3.3-2}

for all $a \in B(b, \frac{2}{i})$. Note that $E(c, T, i)$ is a Borel set since Df is Borel measurable. Note that from (3.3.1) and (3.3.2) follows the estimate

$$\frac{1}{t}|T(a - b)| \leq |f(a) - f(b)| \leq t|T(a - b)| \tag{3.3.3}$$

{eq3.3-3}

holding for all $b \in E(c, T, i)$ and $a \in B(b, \frac{2}{i})$.

(iii). We next show that if $b \in E(c, T, i)$, then

$$\left(\frac{1}{t} + \epsilon\right)^n |\det T| \leq Jf(b) \leq (t - \epsilon)^n |\det T|.$$

To see this, first note that Df is a linear map. Thus there exists an orthogonal map $O : \mathbb{R}^n \to \mathbb{R}^m$ and a symmetric map $S : \mathbb{R}^n \to \mathbb{R}^n$ (cf. (3.2.2)) such that $Df = O \circ S$. Then

$$Jf(b) = [\![Df(b)]\!] = |\det S|.$$

By (3.3.1),

$$\left(\frac{1}{t} + \epsilon\right) |Tv| \leq |(O \circ S)v| = |Sv| \leq (t - \epsilon)|Tv|$$

for all $v \in \mathbb{R}^n$, and so

$$\left(\frac{1}{t} + \epsilon\right) |v| \leq |(S \circ T^{-1})v| \leq (t - \epsilon)|v|$$

for all $v \in \mathbb{R}^n$. Thus

$$(S \circ T^{-1})(B(0, 1)) \subset B(0, t - \epsilon),$$

so that

$$|\det(S \circ T^{-1})|\alpha(n) \leq \mathcal{L}^n(B(0, t - \epsilon)) = \alpha(n)(t - \epsilon)^n,$$

and hence

$$|\det S| \leq (t - \epsilon)^n |\det T|.$$

The proof of the reverse inequality follows from the fact that

$$|(S \circ T^{-1})v| \geq \left(\frac{1}{t} + \epsilon\right),$$

and thus

$$B\left(0, \frac{1}{t} + \epsilon\right) \subset (S \circ T^{-1})(B(0, 1)).$$

(iv). Relabel the countable collection $\{E(c, T, i) : c \in C, T \in S, i \in \mathbb{N}\}$ as $\{E_k\}_{k=1}^{+\infty}$. Choose any $b \in B$, write $Df = O \circ S$, and choose $T \in S$ such that

$$\operatorname{Lip}(T \circ S^{-1}) \leq \left(\frac{1}{t} + \epsilon\right)^{-1}, \quad \operatorname{Lip}(S \circ T^{-1}) \leq t - \epsilon.$$

Now choose $i \in \mathbb{N}$ and $c \in C$ such that $|b - c| < \frac{1}{i}$ and

$$|f(a) - f(b) - Df(b) \cdot (a - b)| \leq \frac{\epsilon}{\mathrm{Lip}(T^{-1})}|a - b| \leq \epsilon|T(a - b)|$$

for all $a \in B(b, \frac{2}{i})$. Then by (iii), $b \in E(c, T, i)$. Since this holds for all $b \in B$, this proves assertion (i).

(v). Next choose any set $E_k = E(c, T, i)$. Let $T_k := T$. By (3.3.3),

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

for all $b \in E_k$, $a \in B(b, \frac{2}{i})$. Since $E_k \subset B(c, \frac{1}{i}) \subset B(b, \frac{2}{i})$, we have

$$\frac{1}{t}|T_k(a - b)| \leq |f(a) - f(b)| \leq t|T_k(a - b)| \tag{3.3.4}$$ {eq3.3-4}

holding for all $a, b \in E_k$. Thus $f|_{E_k}$ is one–to–one.

(vi). Finally notice that (3.3.4) implies

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

Thus (iii) provides the esitmate

$$t^{-n}|\det T_k| \leq Jf|_{E_k} \leq t^n|\det T_k|,$$

which proves assertion (iii). The proof is complete. □

. . . .

3.3.2. *Proof of the Area Formula.*

t3.3-1 **Theorem 3.3.1** (The Area Formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \leq m$. Then for each $\mathcal{L}^n$−measurable subset $A \subseteq \mathbb{R}^n$,*

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

Proof.

(i). In view of Rademacher's Theorem (cf. (3.1.2)), we may assume that $Df(x)$ and $Jf(x)$ exist for all $x \in A$. We may also assume that $\mathcal{L}^n(A) < +\infty$, for otherwise both sides of the equality are $+\infty$.

(ii). Suppose now that $A \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$. Fix $t > 1$ and choose Borel sets $\{E_k\}_{k=1}^{+\infty}$ as in Lemma (3.3.3). That is,

1. $B = \cup_{k=1}^{+\infty} E_k$,
2. $f|_{E_k}$ is one–to–one, $k = 1, 2, \ldots,$
3. For each $k = 1, 2, \ldots,$ there exists a symmetric automorphism $T_k : \mathbb{R}^n \to \mathbb{R}^n$ such that

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

and

$$t^{-n}|\det T_k| \leq Jf|_{E_k} \leq t^n|\det T_k|.$$

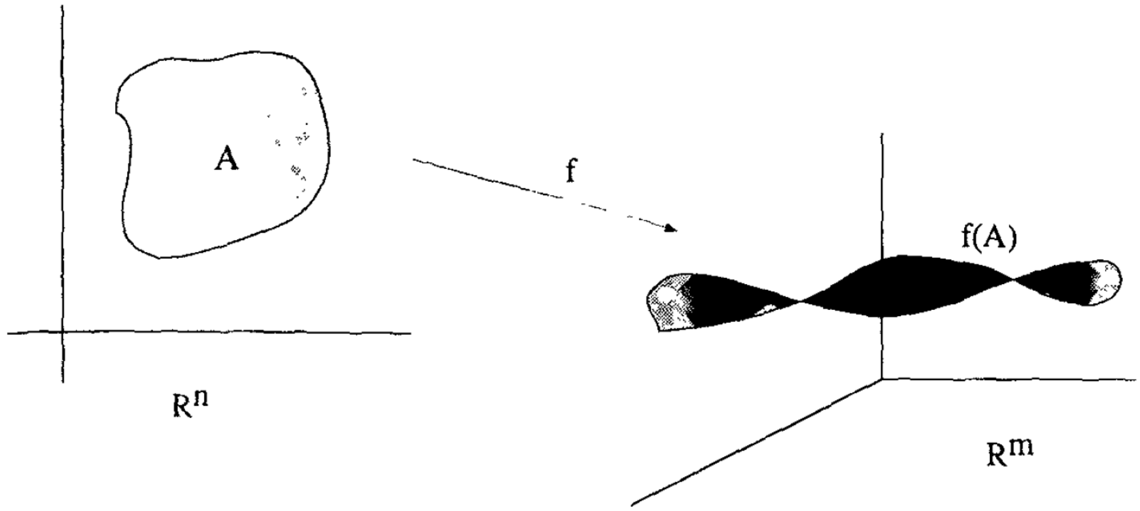

FIGURE 3.3.1. The Area Formula.

Upon passing to the collection $F_k := E_k \setminus (\cup_{i=1}^{k-1} E_k)$ if necessary, we may also suppose that the set $\{E_k\}_{k=1}^{+\infty}$ are disjoint. Define $\mathcal{B}_k$ as in the proof of Lemma (3.3.2), that is,

$$\mathcal{B}_k := \{Q : Q = (a_1, b_1] \times \cdots \times (a_n, b_n], a_i := \frac{c_i}{k}, b_i := \frac{c_i + 1}{k}, c_i \in \mathbb{Z}, i = 1, \dots, n\}.$$

Set

$$F_j^i := E_j \cap Q_i \cap A, \quad Q_i \in \mathcal{B}_k, \quad j = 1, \dots, n.$$

Then the sets F_j^i are disjoint because $\{E_k\}_{k=1}^{+\infty}$ is disjoint, and $A = \cup_{i,j=1}^{+\infty} F_j^i$.

(iii). We claim that

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

To see this, put

$$g_k := \sum_{i,j=1}^{+\infty} \mathbb{1}_{f(F_j^i)}.$$

Note that $g_k(y)$ is equal to the number of sets $\{F_j^i\}$ such that $F_j^i \cap f^{-1}(y) \neq \emptyset$. Then $g_k(y) \to \mathcal{H}^0(A \cap f^{-1}(y))$ as $k \to +\infty$. Notice that this is also an increasing sequence. Thus by the Monotone Convergence Theorem,

$$\begin{aligned}
\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}(y)) d\mathcal{H}^n(y) &= \int_{\mathbb{R}^m} \lim_{k \to +\infty} g_k(y)\, d\mathcal{H}^n(y) \\
&\overset{MCT}{=} \lim_{k \to +\infty} \int_{\mathbb{R}^m} g_k(y)\, d\mathcal{H}^n(y) \\
&= \lim_{k \to +\infty} \sum_{i,j=1}^{+\infty} \mathcal{H}^n(f(F_j^i)),
\end{aligned}$$

where the last inequality follows from the fact that $\{F_j^i\}$ is disjoint.

(iv). Next note that

$$\mathcal{H}^n(f(F_j^i)) = \mathcal{H}^n(f|_{E_j}(F_j^i)) = \mathcal{H}^n(f|_{E_j} \circ T_j^{-1} \circ T_j(F_j^i)) \le t^n \mathcal{L}^n(T_j(F_j^i))$$

and

$$\mathcal{L}^n(T_j(F_j^i)) = \mathcal{H}^n(T_j \circ (f|_{E_j})^{-1} \circ f|_{E_j}(F_j^i)) \le t^n \mathcal{H}^n(f(F_j^i))$$

by Lemma (3.3.3) (cf. (2.4.1)). Thus

$$\begin{aligned} t^{-2n}\mathcal{H}^n(f(F_j^i)) &\le t^{-n}\mathcal{L}^n(T_j(F_j^i)) \\ &= t^{-n}|\det T_j|\mathcal{L}^n(F_j^i) \\ &\le \int_{F_j^i} Jf(x)\, d\mathcal{L}^n(x) \\ &\le t^n|\det T_j|\mathcal{L}^n(F_j^i) \\ &= t^n\mathcal{L}^n(T_j(F_j^i)) \\ &\le t^{2n}\mathcal{H}^n(f(F_j^i)) \end{aligned}$$

(cf. Lemmas (3.3.1) and (3.3.3)). Now summing on i and j, and recalling that $A = \cup_{i,j=1}^{+\infty} F_j^i$, we have

$$t^{-2n}\sum_{i,j=1}^{+\infty}\mathcal{H}^n(f(F_j^i)) \le \int_A Jf(x)\, d\mathcal{L}^n(x) \le t^{2n}\sum_{i,j=1}^{+\infty}\mathcal{H}^n(f(F_j^i)).$$

Letting $k \to +\infty$, we have by (iii) that

$$t^{-2n}\int_{\mathbb{R}^m}\mathcal{H}^0(A\cap f^{-1}(y))\, d\mathcal{H}^n(y) \le \int_A Jf(x)\, d\mathcal{L}^n(x) \le t^{2n}\int_{\mathbb{R}^m}\mathcal{H}^0(A\cap f^{-1}(y))\, d\mathcal{H}^n(y).$$

Finally, taking the limit as $t \to 1^+$ shows that

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

which completes the proof for the case $A \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$.

(v). Now consider the case $A \subset \{x \in \mathbb{R}^n : Jf(x) = 0\}$. Fix $\epsilon > 0$. We factor $f := p \circ g$, where

$$g : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n, \quad g(x) := (f(x), \epsilon x), \quad x \in \mathbb{R}^n,$$

and

$$p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m, \quad p(y,z) := y, \quad y \in \mathbb{R}^m,\ z \in \mathbb{R}^n.$$

(vi). We now claim that there exists a constant $C > 0$ such that

$$0 < Jg(x) \le C\epsilon$$

for all $x \in A$. To prove this claim, write $g = (f^1, \dots, f^m, \epsilon x_1, \dots, \epsilon x_m)$. Then

$$Dg(x) = \begin{bmatrix} Df(x) \\ \epsilon I \end{bmatrix}.$$

Since $Jg(x)^2$ equals the sum of squares of the $(n\times n)$ subdeterminants of $Dg(x)$ according to the Binet–Cauchy Formula (cf. (3.2.4)), we see that

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

Moreover, since $|Df| \leq \operatorname{Lip}(f) < +\infty$, we may use the Binet–Cauchy formula to also compute

$$Jg(x)^2 = Jf(x)^2 + \{\text{sum of squares of terms each involving at least one } \epsilon\} \leq C\epsilon^2$$

for each $x \in A$.

(vii). Since $p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ is a projection, $\operatorname{Lip}(p) \leq 1$, and we can compute using the first case $A \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$

$$\begin{aligned}\mathcal{H}^n(f(A)) &\leq \mathcal{H}^n(g(A)) \\ &\leq \int_{\mathbb{R}^{n+m}} \mathcal{H}^0(A \cap g^{-1}(y,z))\, d\mathcal{H}^n(y,z) \\ &= \int_A Jg(x)\, d\mathcal{L}^n(x) \\ &\leq C\epsilon\mathcal{L}^n(A).\end{aligned}$$

Letting $\epsilon \to 0$, we conclude that $\mathcal{H}^n(f(A)) = 0$, and thus

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

since $\operatorname{supp} \mathcal{H}^0(A \cap f^{-1}(y)) \subset f(A)$. But then since $Jf(x) = 0$ on A by the assumption, it follows

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

as required.

(viii). In the general case, write $A := A_1 \cup A_2$, with $A_1 \subset \{x \in \mathbb{R}^n : Jf(x) > 0\}$, $A_2 \subset \{x \in \mathbb{R}^n : Jf(x) = 0\}$, and apply the above arguments. The proof is complete. □

. . . .

3.3.3. *Change of Variables Formula.*

t3.3-2 **Theorem 3.3.2.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz, $n \leq m$. Then for each $\mathcal{L}^n$−integrable function $g : \mathbb{R}^n \to \mathbb{R}$,*

$$\int_{\mathbb{R}^n} g(x)Jf(x)\, dx = \int_{\mathbb{R}^m} \left[\sum_{x \in f^{-1}(y)} g(x)\right] d\mathcal{H}^n(y).$$

Proof.

(i). Consider first the case $g \geq 0$. Recall that the sequence $\{s_n\}_{n=1}^{+\infty}$ of simple functions defined by

$$s_n(x) := \sum_{k=1}^{n2^n} \frac{k}{2^n} \mathbb{1}_{g^{-1}[\frac{k}{2^n}, \frac{k+1}{2^n})}(x) + n\mathbb{1}_{g^{-1}[n,+\infty]}(x)$$

satisfies $s_n \to g$ and

$$0 \leq g_1 \leq g_2 \leq \cdots .$$

Thus the Monotone Convergence Theorem implies that

$$\int_{\mathbb{R}^n} g(x)Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^n} \lim_{n\to+\infty} s_n(x)Jf(x)\, d\mathcal{L}^n(x)$$

$$
\begin{aligned}
&\overset{MCT}{=} \int_{\mathbb{R}^n} \lim_{n\to+\infty} s_n(x) Jf(x)\, d\mathcal{L}^n(x) \\
&\overset{B.L.}{=} \sum_{k=1}^{+\infty} \frac{k}{2^n} \int_{g^{-1}[\frac{k}{2^n},\frac{k+1}{2^n})} Jf(x)\, d\mathcal{L}^n(x) \\
&= \sum_{k=1}^{+\infty} \frac{k}{2^n} \int_{\mathbb{R}^m} \mathcal{H}^0 \left(g^{-1}\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \cap f^{-1}(y) \right) d\mathcal{H}^n(y) \\
&\overset{B.L.}{=} \int_{\mathbb{R}^m} \sum_{n=1}^{+\infty} \frac{k}{2^n} \sum_{x\in f^{-1}(y)} \mathbb{1}_{g^{-1}[\frac{k}{2^n},\frac{k+1}{2^n})}(x)\, d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^m} \sum_{x\in f^{-1}(y)} \sum_{n=1}^{+\infty} \frac{k}{2^n} \mathbb{1}_{g^{-1}[\frac{k}{2^n},\frac{k+1}{2^n})}(x)\, d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^m} \left[\sum_{x\in f^{-1}(y)} g(x) \right] d\mathcal{H}^n(y).
\end{aligned}
$$

(ii). Now in the case that g is any $\mathcal{L}^n-$integrable function, write $g = g^+ - g^-$ and apply the above case (i). The proof is complete. □

. . . .

3.3.4. *Applications.*

Example 3.3.1 (Length of a Curve ($n = 1, m \geq 1$)). *Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz and one–to–one. Write*

$$f = (f^1, \dots, f^m), \quad Df = (\dot{f}^1, \dots, \dot{f}^n),$$

so that

$$Jf = |Df| = |\dot{f}|.$$

For any $-\infty < a < b < +\infty$, define the curve

$$C := f([a,b]) \subset \mathbb{R}^m.$$

Then by the Area Formula

$$
\begin{aligned}
\int_a^b |\dot{f}(t)|\, dt &= \int_{[a,b]} Jf(x)\, d\mathcal{L}^1(x) \\
&= \int_{\mathbb{R}^m} \mathcal{H}^0([a,b] \cap f^{-1}(y))\, d\mathcal{L}^1(y) \\
&= \mathcal{H}^1(C).
\end{aligned}
$$

Example 3.3.2 (Surface Area of a Graph ($n \geq 1, m = n+1$)). *Assume that $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz and define $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by*

$$f(x) := (x, g(x)).$$

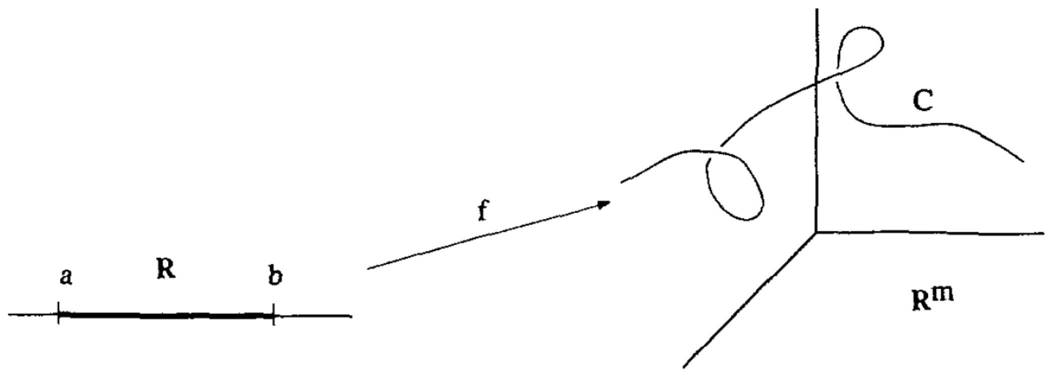

FIGURE 3.3.2. Length of a Curve.

Note that $f = \Gamma(g)$. Then

$$Df(x) = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \\ \frac{\partial}{\partial x_1} g(x) & \cdots & \frac{\partial}{\partial x_n} g(x) \end{bmatrix}.$$

By the Binet–Cauchy formula,

$$(Jf)^2 = \text{sum of squares of } n \times n \text{ subdeterminants} = 1 + |Dg|^2,$$

so that $Jf = (1 + |Dg|^2)^{1/2}$. Now for each open set $\Omega \subset \mathbb{R}^n$, recall the graph of g over Ω :

$$\Gamma(g, \Omega) = \{(x, f(x)) : x \in \Omega\} \subset \mathbb{R}^{n+1}.$$

Then by the Area Formula

$$\begin{aligned} \int_\Omega (1 + |Dg(x)|^2)^{1/2} \, d\mathcal{L}^n(x) &= \int_\Omega Jf(x) \, d\mathcal{L}^n(x) \\ &= \int_{\mathbb{R}^{n+1}} \mathcal{H}^0(\Omega \cap f^{-1}(y)) \, d\mathcal{H}^n(y) \\ &= \mathcal{H}^n(\Omega). \end{aligned}$$

Example 3.3.3 (Surface Area of a Parametric Hypersurface ($n \geq 1$, $m = n + 1$)). *Suppose that $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ is one–to–one and Lipschitz. Write*

$$f = (f^1, \ldots, f^{n+1})$$

and

$$Df(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f^1(x) & \cdots & \frac{\partial}{\partial x_n} f^1(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f^{n+1}(x) & \cdots & \frac{\partial}{\partial x_n} f^{n+1}(x) \end{bmatrix}.$$

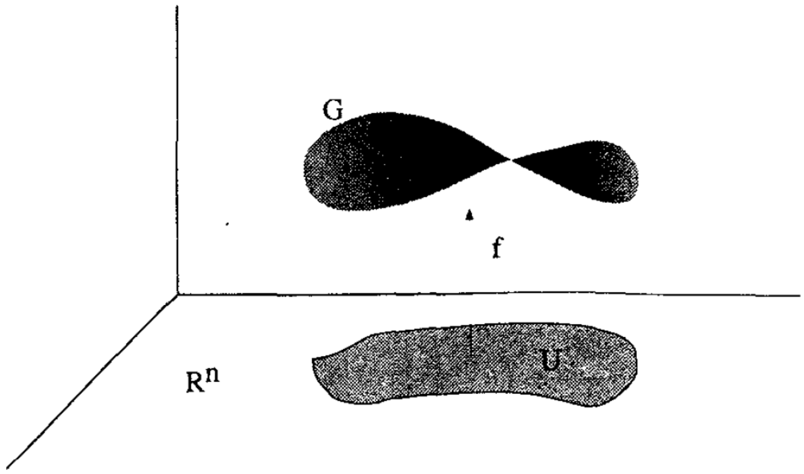

FIGURE 3.3.3. Surface Area of a Graph.

Then by the Binet–Cauchy formula,

$$
\begin{aligned}
(Jf)^2 &= \textit{sum of squares of } n \times n \textit{subdeterminants} \\
&= \sum_{k=1}^{n+1} \left[\frac{\partial(f^1, \ldots, f^{k-1}, f^{k+1}, \ldots, f^{n+1})}{\partial x_1, \ldots, x_n} \right]^2,
\end{aligned}
$$

where

$$
\frac{\partial(f^1, \ldots, f^{k-1}, f^{k+1}, \ldots, f^{n+1})}{\partial x_1, \ldots, x_n}
$$

denotes the Jacobian of the function with gradient matrix

$$
\begin{bmatrix}
\frac{\partial}{\partial x_1} f^1(x) & \cdots & \frac{\partial}{\partial x_n} f^1(x) \\
\vdots & & \vdots \\
\frac{\partial}{\partial x_1} f^{k-1}(x) & \cdots & \frac{\partial}{\partial x_n} f^{k-1}(x) \\
\frac{\partial}{\partial x_1} f^{k+1}(x) & \cdots & \frac{\partial}{\partial x_n} f^{k+1}(x) \\
\vdots & & \vdots \\
\frac{\partial}{\partial x_1} f^{n+1}(x) & \cdots & \frac{\partial}{\partial x_n} f^{n+1}(x)
\end{bmatrix}.
$$

For each open set $\Omega \subset \mathbb{R}^n$, write

$$
S := f(\Omega) \subset \mathbb{R}^{n+1}.
$$

Then by the Area Formula

$$\begin{aligned}\int_\Omega \left(\sum_{k=1}^{n+1} \left[\frac{\partial(f^1, \ldots, f^{k-1}, f^{k+1}, \ldots, f^{n+1})}{\partial x_1, \ldots, x_n} \right]^2 \right)^{\frac{1}{2}} \, d\mathcal{L}^n(x) &= \int_\Omega Jf(x)\, d\mathcal{L}^n(x) \\ &= \int_{\mathbb{R}^{n+1}} \mathcal{H}^0(\Omega \cap f^{-1}(y))\, d\mathcal{H}^n(y) \\ &= \mathcal{H}^n(S).\end{aligned}$$

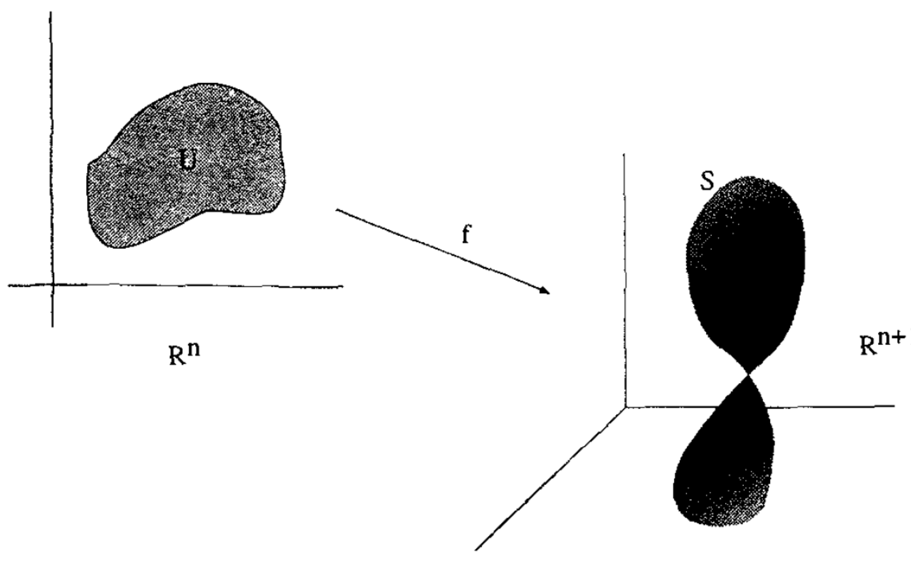

FIGURE 3.3.4. Surface Area of a Parametric Hypersurface.

Example 3.3.4 (Submanifolds). *Let $M \subset \mathbb{R}^m$ be a Lipschitz $n-$dimensional embedded submanifold. Suppose that $\Omega \subset \mathbb{R}^n$ and let $f : \Omega \to M$ be coordinates for M. Let $A \subset f(\Omega)$. Let $A \subset f(\Omega) \subset M$, A Borel, and let $B := f^{-1}(A) \subset \Omega$. Define the metric $g : M \to \mathbb{R}$ on M by*

$$g_{ij} = g\left(\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j} \right) := \frac{\partial f}{\partial x_i} \cdot \frac{\partial f}{\partial x_j}, \quad i, j = 1, \ldots, n,$$

and

$$g := \det((g_{ij})_{n\times n}).$$

Then

$$Df \circ (Df)^* = (g_{ij})_{n\times n},$$

and so

$$Jf = (\det(Df \circ (Df)^*))^{\frac{1}{2}} = g^{\frac{1}{2}}.$$

Thus by the Area Formula,

$$\begin{aligned}\int_B g^{\frac{1}{2}}\, d\mathcal{L}^n(x) &= \int_B Jf(x)\, d\mathcal{L}^n(x) \\ &= \int_{\mathbb{R}^m} \mathcal{H}^0(B \cap f^{-1}(y))\, d\mathcal{H}^n(y)\end{aligned}$$

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

Here $\mathcal{H}^n(A)$ represents the "volume" of A in M.

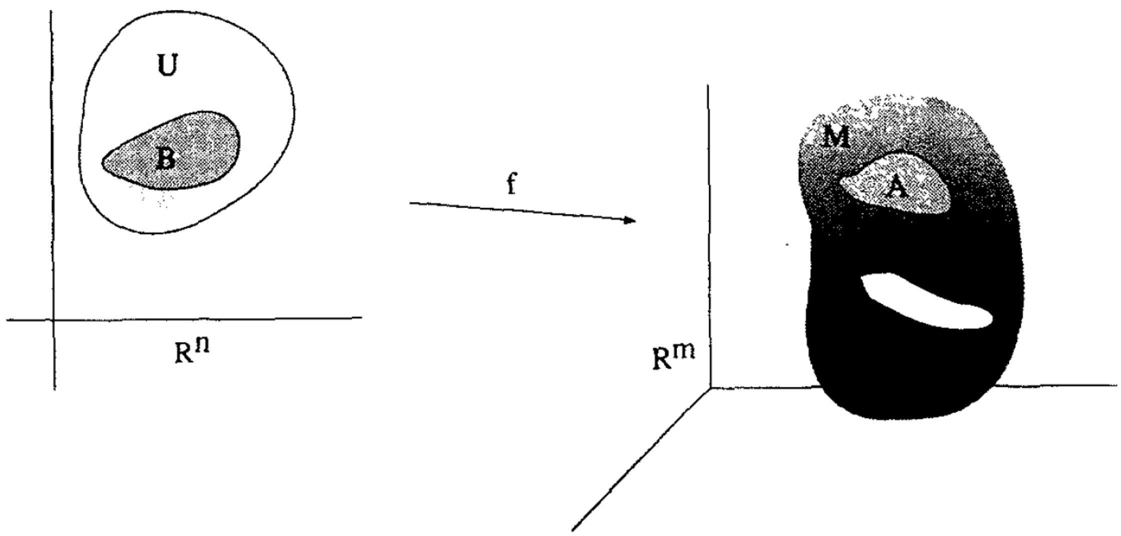

FIGURE 3.3.5. Volume of a Submanifold.

3.4. **The Coarea Formula.**

References

1. Lawrence C. Evans and Ronald F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. MR 1158660