THE AREA FORMULA FOR THE HAUSDORFF MEASURE

VANESSA PIERRE NGHAYWI

Thesis Advisor: Dr. Jessica Merhej

Notre Dame University, Louaize

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Mathematics in the Department of Mathematics and Statistics in the Faculty of Natural and Applied Sciences

June 27, 2018

The Area Formula For The Hausdorff Measure

Vanessa Pierre Nghaywi

Approved by:

Dr. Jessica Merhej

Dr. Bassem Ghalayini

Dr. Joseph Melkoune

Abstract

In this thesis, we will prove a very important theorem in real analysis called The Area formula for the Hausdorff Measure. This theorem is an extension of the well known theorem : the Change of variables formula for the Lebesgue measure. In this thesis, we will define the Hausdorff measure and prove some of its properties. We will also define Lipschitz functions and prove some of its properties also. Then, we continue the thesis by proving all the lemmas needed to finalize the proof of the Area formula for the Hausdorff measure. Finally, we finish this thesis by showing three applications of the Area formula.

Acknowledgments

I must express my very profound gratitude to my parents and to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

I am heartily thankful to my advisor Dr. Jessica Merhej whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Her guidance and immense knowledge helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my master study.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Bassem Ghalayini and Dr. Joseph Melkoune for their insightful comments and encouragement during my thesis defense.

Contents

Chapter 1

Notations

$\mathcal{L}^1$	The 1-dimensional Lebesgue measure
$\mathcal{L}^n$	The n-dimensional Lebesgue measure
O_+	Positive functions
$f\restriction_E$	f restricted to the set E
a.e.	almost everywhere
λ^*	The Lebesgue outer measure
$\mathcal{H}^s$	s-dimensional Hausdorff measure
$D_\mu v$	the derivative of v with respect to μ
e_i	$(0,0,\cdots,1,0,\cdots)$ with 1 in the ith slot
$x=(x_1,\cdots,x_n)$	a typical point in $\mathbb{R}^n$
$B(x,r)$	$\{y\in\mathbb{R}^n,\ \lvert x-y\rvert\le r\}$ = closed ball with center x, radius r}
$\alpha(s)$	$\dfrac{\pi^{\frac{s}{2}}}{\gamma\left(\frac{s}{2}+1\right)}$ $(0\le s<\infty)$
$\alpha(n)$	volume of the unit ball in $\mathbb{R}^n$
χ_A	indicator function of the set A
$\bar{A}$	closure of the set A
$S_a(A)$	Steiner Symmetrization of the set A
$\bar{f}$	an extension of f
Df	derivative of f
$[Df]$	measure of the gradient of f
Jf	Jacobian of f
$Lip(f)$	Lipschitz constant of f
$\nu\ll\mu$	ν is absolutely continuous with respect to μ
$[[L]]$	jacobian of a linear map L
$\Lambda(m,n)$	$\left\{\lambda:\{1,\cdots,n\}\to\{1,\cdots,m\};\lambda\,is\,increasing\right\}$
A^0	interior of A
$x.y$	$x_1y_1+\cdots+x_ny_n$

Chapter 2

Introduction

In measure theory, the Lebesgue measure, named after the french Mathematicien Henri Lebesgue is the standard way of assigning a measure to subsets of n-dimensional euclidean space. For $n = 1$, the lebesgue measure coincides with measuring the length; for $n = 2$, it coincides with measuring the area; and for $n = 3$, it coincides with measuring the volume and so on. For instance, the Lebesgue measure of the interval [0, 1] in the real numbers is its length in the everyday sense of the word, specifically, 1. For the general case, that is in $\mathbb{R}^n$ the Lebesgue measure is called the n-dimensional volume, n-volume, or simply volume. It is used throughout real analysis, in particular to define Lebesgue integration. Sets that can be assigned a Lebesgue measure are called Lebesgue measurable. Henri Lebesgue described this measure in the year 1901, followed the next year by his description of the Lebesgue integral. Both were published as part of his dissertation in 1902.
Now we will start by defining the Lebesgue outer measure on any set $A \subset \mathbb{R}^n$, and then the Lebesgue measure.

Definition 2.0.1. If $B = I_1 \times I_2 \times \cdots \times I_n$ where $I_n = [a_n, b_n]$ are intervals, then the volume of B is defined to be :

$$V(B) = (b_1 - a_1) \times \cdots \times (b_n - a_n).$$

For any subset A of $\mathbb{R}^n$, we can define the outer measure of A by

$$\lambda^* = \inf \Big\{ \sum_{B \in \mathcal{C}} V(B)\, ;\ \mathcal{C} \text{ is a countable collection of boxes whose union cover } A \Big\}.$$

We then define the set A to be Lebesgue measurable, if for every set $E \in \mathbb{R}^n$ we have

$$\lambda^*(E) = \lambda^*(A \cap E) + \lambda^*(E \cap A^c).$$

These Lebesgue measurable sets form a σ-algebra, and the Lebesgue measure is defined by $\mathcal{L}^n(A) = \lambda^*(A)$ for any Lebesgue measurable set A.

The importance of the Lebesgue measure comes from the fact that we can find the area between a Lebesgue measurable function and a measurable set, which is also known as the Lebesgue integral. The Lebesgue integral plays an important role in probability theory, real

analysis, and many other fields in the mathematical sciences such as differential geometry. Since manifolds act locally like $\mathbb{R}^n$, we can find ways to define integration on manifolds using the Lebesgue measure or an equivalent (See book [2]).
A very important application of the Lebesgue measure is the Change of Variables formula .

Theorem 2.0.2. *Change of Variables for $\mathcal{L}^n$.*
Let $U \subset V \subset \mathbb{R}^n$. U is lebesgue measurable and V is open.
Let $T : V \longrightarrow \mathbb{R}^n$ be a continuous and one-to-one function on U.

$$T'(U) \text{ exists for all } u \in U \text{ and } \mathcal{L}^n(T(v-u)) = 0.$$

Then,

$$\int_{T(U)} f\, d\mathcal{L}^n = \int_U (f \circ T)\, |J_T|\, d\mathcal{L}^n \text{ for all } f \in O_+.$$

Notice that $X = T(U)$. In other words $\int_X f(x)\, dx = \int_U f(Tu)\, |J_T(u)|\, du$ *for all* $f \in O_+$.
Where J_T is the jacobian of T.
$(\mathcal{L}(T(U)) = |J_T| \mathcal{L}(U))$.

Its importance come from the fact that it connects the area of a surface to its area under a certain transformation.
However, the disadvantage of the Lebesgue measure is that it only can measure the n-dimensional volume of n-dimensional spaces. That is the n-dimensional Lebesgue measure does not see the difference between lesser dimensional objects. For example $\mathcal{L}^3$ does not see the difference between a one dimensional line and a two dimensional plane; both have a Lebesgue measure zero.
So mathematicians needed to introduce a new measure which is an extension of the Lebesgue measure , but instead it can give the area of an object according to its dimension even if it lives in a bigger dimensional space. For example if we have a 2-dimensional surface living in $\mathbb{R}^5$, we need a measure that gives us the area of this surface even if it is not living in $\mathbb{R}^2$. This new measure is known as the Hausdorff measure and it was introduced in 1918 by the mathematician Felix Hausdorff. We will see throughout this thesis that the zero dimensional Hausdorff measure is just the counting measure, that is, the number of points in the set (if the set is finite) or ∞ if the set is infinite. The one-dimensional Hausdorff measure of a simple curve in $\mathbb{R}^n$ is equal to the length of the curve. Likewise, the two dimensional Hausdorff measure of a measurable subset of $\mathbb{R}^n$ is proportional to the area of the set.
Thus, the concept of the Hausdorff measure generalizes counting, length, area and volume like the Lebesgue measure; the only difference is that the Hausdorff measure can measure the length , area and volume of $1, 2$ and 3 dimensional objects that live in a higher dimensional space.
Now we will give the mathematical definition of the Hausdorff measure.

Definition 2.0.3. 1. Let $A \subset \mathbb{R}^n$, $0 \leq s < \infty$, $0 \leq \delta < \infty$.
Let us define

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

and where $\alpha(s) = \dfrac{\pi^{\frac{s}{2}}}{\gamma\left(\frac{s}{2}+1\right)}$.

2. For $A \subset \mathbb{R}^n$ and $0 \leq s < \infty$, let us define

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

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

In order to show that $\mathcal{H}^s$ is well defined, we show that $\mathcal{H}^s_\delta$ increases as δ decreases. So let $A \subset \mathbb{R}^n$, and $\delta_2 < \delta_1$. Notice that

$$\left\{\{C_j\}_{j=1}^{\infty} ; A \subset \bigcup_{j=1}^{\infty} C_j \text{ ; diam } C_j \leq \delta_2\right\} \subseteq \left\{\{C_j\}_{j=1}^{\infty} ; A \subset \bigcup_{j=1}^{\infty} C_j \text{ ; diam } C_j \leq \delta_1\right\}.$$

And thus,

$$\inf\left\{\bigcup_{j=1}^{\infty} C_j \text{ , diam } C_j \leq \delta_1\right\} \leq \inf\left\{\bigcup_{j=1}^{\infty} C_j \text{ , diam } C_j \leq \delta_2\right\}.$$

This implies that $\mathcal{H}^s_{\delta_1} \leq \mathcal{H}^s_{\delta_2}$. Hence, if δ decreases, $\mathcal{H}^s_\delta$ increases. So $\lim\limits_{\delta \to 0} \mathcal{H}^s_\delta$ exists and $\lim\limits_{\delta \to 0} \mathcal{H}^s_\delta = \sup\limits_{\delta > 0} \mathcal{H}^s_\delta$. Notice that the Hausdorff measure is computed in terms of arbitrary coverings of small diameters whereas the Lebesgue measure is computed in terms of coverings by cubes.

The purpose of this thesis is to prove a far reaching generalization of the change of variables formula called the Area formula of the Hausdorff measure. In order to establish this big theorem, we first need to prove some properties of Hausdorff measure. A very important theorem called "The isodiametric inequality " will be handled , which states that the n-dimensional Lebesgue measure is equal to the n- dimensional Hausdorff measure on an n-dimensional space. This shows that the Hausdorff measure and the Lebesgue measure coincides on $\mathbb{R}^n$.

We proceed by defining Lipschitz functions which by themselves are a generelization of differentiable functions and all its properties as well as linear maps and Jacobians. A very important theorem will arise in this section: Rademachers theorem. This theorem states that any locally Lipschitz function f mapping from a lower dimensional space onto a higher dimensional space is differentiable $\mathcal{L}^n$ almost everywhere.
Then, we will build up the math by handling several big Lemmas to get to the Area formula , which is the same idea as the area formula of the Lebesgue measure but now upgraded to the Hausdorff measure and we will be integrating against Lipschitz functions. Finally we will apply the area formula on 3 examples , to finish our thesis.

Chapter 3

Preliminaries

Definition 3.0.1. (see Section 1.6.2 on page 37 in [1].)
Let μ and v be radon measures on $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, define

$$\overline{D}_\mu v(x) = \begin{cases} \limsup_{r\to 0} \frac{v(B(x,r))}{\mu(B(x,r))} & if\ \mu(B(x,r)) > 0\ \forall r > 0 \\ +\infty & if\ \mu(B(x,r)) = 0\ for\, some\, r > 0 \end{cases}$$

$$\underline{D}_\mu v(x) = \begin{cases} \liminf_{r\to 0} \frac{v(B(x,r))}{\mu(B(x,r))} & if\ \mu(B(x,r)) > 0\ \forall r > 0 \\ +\infty & if\ \mu(B(x,r)) = 0\ for\, some\, r > 0 \end{cases}$$

If $\overline{D}_\mu v(x) = \underline{D}_\mu v(x) < \infty$, we say that v is differentiable with respect to μ at x and write

$$D_\mu v(x) = \overline{D}_\mu v(x) = \underline{D}_\mu v(x)$$

where $D_\mu v$ is the derivative of v with respect to μ.

Definition 3.0.2. Absolute continuity (see Section 1.6.2 on page 40 in [1].)
The measure ν is absolutely continuous with respect to μ , written $\nu \ll \mu$, provided $\mu(A) = 0$ implies that $\nu(A) = 0$ for all $A \subset \mathbb{R}^n$.

Theorem 3.0.3. *Every Lipschitz function is absolutely continuous .*

Proof. Let $g : [a,b] \longrightarrow \mathbb{R}$ be a lipschitz function, then $|g(b) - g(a)| \leq C|b-a|$, for some $C \in \mathbb{R}^n$. Fix $\epsilon > 0$ and let $P = \{[a_i, b_i]\}_{i=1}^n$ be a partition of $[a,b]$ such that $\sum_{i=1}^{n} |b_i - a_i| < \frac{\epsilon}{C}$, then

$$\begin{aligned} \sum_{i=1}^{n} |f(b_i) - f(a_i)| &\leq C \sum_{i=1}^{n} |b_i - a_i| \\ &\leq C\frac{\epsilon}{C} \\ &= \epsilon. \end{aligned}$$

□

Theorem 3.0.4. *Caratheodory's Criterion (see Theorem 5 page 9 in [1].)*
Let μ be a measure on $\mathbb{R}^n$. Suppose that $\mu(A \cup B) = \mu(A) + \mu(B)$ for all sets A, B in $\mathbb{R}^n$ such that dist$(A, B) > 0$. Then, μ is a Borel Measure.

Theorem 3.0.5. *Fubini's Theorem (see Theorem 2.37 page 67 in [4])*
Suppose that $(X, \mathcal{M}, \mu)$and$(Y, \mathcal{N}, \nu)$ are σ -finite measure spaces.
If $f \in \mathcal{L}^1(\mu \times \nu)$, then $f_x \in \mathcal{L}^1(\nu)$ for a.e. $x \in X$, $f^y \in \mathcal{L}^1(\mu)$ for a.e. $y \in Y$,the a.e.defined functions $g(x) = \int f_x \, d\nu$ and $h(x) = \int f^y \, d\nu$ are in $\mathcal{L}^1(\mu)$ and in $\mathcal{L}^1(\nu)$ respectively and

$$\int f \, d(\mu \times \nu) = \int \left[\int f(x, y) \, d\nu(y) \right] d\mu(x) = \int \left[\int f(x, y) \, d\mu(x) \right] d\nu(y)$$

Definition 3.0.6. n-dimensional Lebesgue measure on $\mathbb{R}^n$ (see Section 1.4 page 26 in [1].)

$$\mathcal{L}^n = \mathcal{L}^{n-1} \times \mathcal{L}^1 \times \mathcal{L}^1 \times \cdots \times \mathcal{L}^1 \text{ n times.}$$

Equivalently $\mathcal{L}^n = \mathcal{L}^{n-k} \times \mathcal{L}^k$ for each $k \in \{1, \cdots, n-1\}$.

Theorem 3.0.7. *Vitali's Covering (see Theorem 1 on page 27 in [1].)*
Let $\mathcal{F}$ be a collection of non degenerate closed balls in $\mathbb{R}^n$ with $\sup\{diam\, B, B \in F\} < \infty$. Then there exists a countable family $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

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

Theorem 3.0.8. *Monotone Convergence Theorem (see book [4])*
Let (X, m, μ) be a measure space.
Let $f, f_1, f_2, \cdots \in O_+$ such that $f_1 \leq f_2 \leq \cdots \leq f$.
If $\lim_{n \to \infty} f_n \to f$ pointwise then

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu.$$

Theorem 3.0.9. *Beppo-Levi (see book [4])*
Let (X, m, μ) be a measure space.
Let $\{f_n\}$ be a sequence in O_+ then,

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} f_n \, d\mu.$$

Theorem 3.0.10. *Dominated Convergence Theorem (see book [4])*
Let (X, m, μ) be a measure space, $f, \{f_n\} \in O_{\mathbb{R}}$ and $\phi \in O_+$.
If :

1. *$f_n \longrightarrow f$, pointwise.*

2. $|f_n| \leq \phi$ *for all* n.

3. $\int \phi' \, d\mu < \infty$, *that is* $\phi \in L_1(\mu)$. *Then,*

$$\int |f_n - f|_{n\to\infty} \, d\mu \longrightarrow 0,$$

and

$$\int f_n \, d\mu \longrightarrow \int f \, d\mu, \, as \, n \to \infty.$$

Chapter 4

Hausdorff Measure

Hausdorff measure
We start this chapter by defining some properties of the Hausdorff Measure.

Theorem 4.0.1. *$\mathcal{H}^s$ is a borel regular measure.*$(0 \leq s < \infty)$.

Proof. We begin by showing $\mathcal{H}^s_\delta$ is a measure, $\forall\ \delta > 0$. Fix $\delta > 0$.

1. Since $\phi \subseteq \phi$ and $diam\ \phi = 0$, then $\mathcal{H}^s_\delta(\phi) \leq \alpha(s)\left(\frac{diam\ \phi}{2}\right)^s = 0$. This implies that $\mathcal{H}^s_\delta(\phi) = 0$.

2. Select sets $\{A_k\}_{k=1}^{\infty} \subset \mathbb{R}^n$ and suppose that each A_k is covered by sets $\{C_j^k\}_{j=1}^{\infty}$ with diam $C_j^k \leq \delta$. Then, $\bigcup_{k=1}^{\infty} A_k$ is covered by $\{C_j^k\}_{k=1}^{\infty}$. Now, using the definition of the Hausdorff measure, we get

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty}\left(\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\ C_j^k}{2}\right)^s\right). \tag{4.0.1}$$

Since (4.0.1) holds for all $C_j^k s$ such that $A_k \subset \bigcup_{j=1}^{\infty} C_j^k$,then by taking the infimum over those C_j^k's, we get

$$\begin{aligned}\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{\infty} A_k\right) &\leq \sum_{k=1}^{\infty}\left(\inf\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\, C_j^k}{2}\right)^s\right)\\ &= \sum_{k=1}^{\infty}\mathcal{H}^s_\delta(A_k),\end{aligned}$$

and hence $\mathcal{H}^s_\delta$ is a measure.

Next we show that $\mathcal{H}^s$ is a measure.

1. Since $\mathcal{H}^s_\delta(\phi) = 0$ for all $\delta > 0$, then taking the supremum over δ, we get that $\sup_{\delta>0} \mathcal{H}^s_\delta(\phi) = \mathcal{H}^s(\phi) = 0$.

2. Fix $\delta > 0$, and as before, select sets $\{A_k\}_{k=1}^\infty \subset \mathbb{R}^n$. Then, since $\mathcal{H}^s_\delta$ is a measure ,

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

Letting $\delta \to 0$ we get, $\mathcal{H}^s\left(\bigcup_{k=1}^\infty A_k\right) \leq \sum_{k=1}^\infty \mathcal{H}^s(A_k)$, finishing the proof that $\mathcal{H}^s$ is a measure.

We proceed by showing that $\mathcal{H}^s$ is a borel measure. To see this, choose sets A , $B \subset \mathbb{R}^n$ such that the distance between A and B is bigger than 0; choose $0 < \delta < \frac{1}{4}\, dist(A, B)$, and suppose that $A \cup B$ is covered by sets $\{C_j\}_{j=1}^\infty$ such that $\operatorname{diam} C_j \leq \delta$. Let us define

$$\begin{aligned} \mathcal{A} &= \{C_j; C_j \cap A \neq \phi\} \\ &\mathit{and} \\ \mathcal{B} &= \{C_j; C_j \cap B \neq \phi\}. \end{aligned}$$

Notice that, $A \subset \bigcup_{C_j \in \mathcal{A}} C_j$, $B \subset \bigcup_{C_j \in \mathcal{B}} C_j$ and $C_j \cap C_i = \phi$ if $C_j \in \mathcal{A}$ and $C_i \in \mathcal{B}$. Thus,

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

This is true for all such $C_j s$ chosen above, hence

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

provided that the distance between A and B is bigger than 4δ and strictly positive. Now letting $\delta \to 0$ we get

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

for all A, B in $\mathbb{R}^n$ such that $\operatorname{dist}(A, B) > 0$. The fact that $\mathcal{H}^s(A \cup B) \leq \mathcal{H}^s(A) + \mathcal{H}^s(B)$ comes from countable subadditivity since we proved that $\mathcal{H}^s$ is a measure. Thus we have that

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

Using Caratheodory's criteria (see theorem (3.0.4)) we get $\mathcal{H}^s$ is a borel measure.

We finish the proof by showing that $\mathcal{H}^s$ is a borel regular measure.
Let's start by noting that $diam\,\overline{C} = diam\,C$ for all $C \subset \mathbb{R}^n$. Hence, we can define $\mathcal{H}^s_\delta$ as

$$\mathcal{H}^s_\delta(A) = \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C_j}{2}\right)^s ; A \subset \bigcup_{j=1}^{\infty} C_j \text{ ; diam}\,C_j \le \delta; C_j \text{ are closed.}\right\}$$

Now, choose $A \subset \mathbb{R}^n$ such that $\mathcal{H}^s(A) < \infty$. Hence $\mathcal{H}^s_\delta(A) < \infty\ \forall\ s > 0$. By the definition of infimum, for each $k \ge 1$,there exist $\{C^k_j\}_{j=1}^{\infty}$, such that $A \subset \bigcup_{j=1}^{\infty} C^k_j, C^k_j$ closed, $\text{diam}\,C^k_j \le \frac{1}{k}$, and,

$$\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C^k_j}{2}\right)^s \le \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C^k_j}{2}\right)^s\right\} + \frac{1}{k}. \tag{4.0.2}$$

which means

$$(4.0.2) \le \mathcal{H}^s_{\frac{1}{k}}(A) + \frac{1}{k}.$$

Let $A_k = \bigcup_{j=1}^{\infty} C^k_j$ and $B = \bigcap_{k=1}^{\infty} A_k$. Notice that B is borel since the C^k_js are closed. Also since $A \subset A_k$ for each k ,we have that $A \subset B$. Furthermore, since $B = \bigcap_{k=1}^{\infty} A_k$, then $B \subset A_k$ for every k, and hence

$$\begin{aligned}
\mathcal{H}^s_{\frac{1}{k}}(B) &\le \mathcal{H}^s_{\frac{1}{k}}(A_k) \\
&= \mathcal{H}^s_{\frac{1}{k}}\left(\bigcup_{j=1}^{\infty} C^k_j\right) \\
&\le \sum_{j=1}^{\infty}\mathcal{H}^s_{\frac{1}{k}}(C^k_j) \\
&\le \sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C^k_j}{2}\right)^s \\
&\le \mathcal{H}^s_{\frac{1}{k}}(A) + \frac{1}{k}
\end{aligned}$$

where the last step comes from (4.0.2). Now if we let $k \to \infty$, we get $\mathcal{H}^s(B) \le \mathcal{H}^s(A)$. The fact that $A \subset B$ gives us the other inequality and hence $\mathcal{H}^s(A) = \mathcal{H}^s(B)$. □

Next, we prove some elementary properties of the Hausdorff measure.

Theorem 4.0.2. *1. $\mathcal{H}^0$ is a counting measure.*

2. $\mathcal{H}^1 = \mathcal{L}^1$ *on* $\mathbb{R}^1$.

3. $\mathcal{H}^s = 0$ *on* $\mathbb{R}^n$ $\forall\, s > n$.

4. $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ $\forall\, \lambda > 0;$ $A \subset \mathbb{R}^n$.

5. $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$ *for each affine isometry* $L : \mathbb{R}^n \longrightarrow \mathbb{R}^n;$ $A \subset \mathbb{R}^n$.

Proof. 1. In order to prove that $\mathcal{H}^0$ is a counting measure, we start by proving that $\mathcal{H}^0(\{a\}) = 1$.

- Let $\delta > 0$. By definition of Hausdorff measure, we have

$$\begin{aligned}\mathcal{H}^0_\delta(\{a\}) &= \inf\left\{\sum_{j=1}^{\infty}\alpha(0)\left(\frac{diam\, C_j}{2}\right)^0,\ \{a\} \subseteq \bigcup_{j=1}^{\infty} C_j, diam\, C_j \le \delta\right\} \\ &= \inf\left\{\sum_{j=1}^{\infty}(diam\, C_j)^0,\ \{a\} \subseteq \bigcup_{j=1}^{\infty} C_j, diam\, C_j \le \delta\right\}.\end{aligned}$$

  Now let $C_1 := B(a,\delta)$. Then $\{a\} \subseteq C_1$ and we get that

$$\mathcal{H}^0_\delta(\{a\}) \le (diam\, C_1)^0 = 1.$$

  To see that $1 \le \mathcal{H}^0_\delta(\{a\})$, take any cover $\{C_j\}_{j=1}^{\infty}$ such that $\{a\} \subset \bigcup_{j=1}^{\infty} C_j$ and $\operatorname{diam} C_j \le \delta$. Then, $\sum_{j=1}^{\infty}(diam\, C_j)^0 \ge 1$. Taking the infimum over such C_j , we get

$$\inf\left\{\sum_{j=1}^{\infty}(diam\, C_j)^0\ ; \{a\} \subseteq \bigcup_{j=1}^{\infty} C_j\ ; diam\, C_j \le \delta\right\} \ge 1$$

  and hence, $\mathcal{H}^0_\delta(\{a\}) \ge 1$.Thus $\mathcal{H}^0_\delta(\{a\}) = 1$ for all δ. Letting $\delta \to 0$ we get that $\mathcal{H}^0(\{a\}) = 1$.

- Next, let us consider countable sets. If $A = \{a_i\}_{i=1}^n$, then by countable additivity $\mathcal{H}^0(\{a_1, \cdots, a_n\}) = \sum_{i=1}^{\infty}\mathcal{H}^0(\{a_i\}) = n$.

- If $A = \{a_i\}_{i=1}^{\infty}$, then also by countable additivity we get

$$\begin{aligned}\mathcal{H}^0(\{a_i\}_{i=1}^{\infty}) &= \sum_{i=1}^{\infty}\mathcal{H}^0(\{a_i\}) \\ &= \infty.\end{aligned}$$

- Finally, if A is uncountable, then there exist $\{a_i\}_{i=1}^{\infty} \subseteq A$ such that $a_i \neq a_j \; \forall \, i \neq j$, and, by countable subadditivity we get

$$\infty = \mathcal{H}^0\left(\{a_i\}_{i=1}^{\infty}\right) \leq \mathcal{H}^0\left(A\right).$$

Thus $\mathcal{H}^0$ is a counting measure.

2. Choose $A \subset \mathbb{R}$ and $\delta > 0$, then

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

For the other inequality, let C_j's be any cover for A so, $A \subseteq \bigcup_{j=1}^{\infty} C_j$. Set $I_k = [k\delta, (k+1)\,\delta]$, for $k \in \mathbb{Z}$. Notice that $diam\,(C_j \cap I_k) \leq \delta$. Using the fact that $\bigcup_{k=-\infty}^{\infty} I_k = \mathbb{R}$, we can see that $\left\{C_j \cap I_k\right\}_{j=1,k=-\infty}^{\infty}$ form a cover for A , since

$$\begin{aligned} A \subseteq \bigcup_{j=1}^{\infty} C_j &= \bigcup_{j=1}^{\infty} (C_j \cap \mathbb{R}) \\ &= \bigcup_{j=1}^{\infty} \left(C_j \cap \bigcup_{k=-\infty}^{\infty} I_k\right) \\ &= \bigcup_{j=1}^{\infty} \left(\bigcup_{k=-\infty}^{\infty} (C_j \cap I_k)\right) \\ &= \bigcup_{j=1,k=-\infty}^{\infty} (C_j \cap I_k). \end{aligned}$$

Also, notice that

$$\begin{aligned} \sum_{k=-\infty}^{\infty} diam\ (C_j \cap I_k) &= diam \left(\bigcup_{k=-\infty}^{\infty} (C_j \cap I_k)\right) \\ &= diam \left(C_j \cap \bigcup_{k=-\infty}^{\infty} I_k\right) \\ &= diam\,(C_j \cap \mathbb{R}) \\ &= diam\ C_j. \end{aligned} \tag{4.0.3}$$

Thus, we have

$$\begin{aligned}
\mathcal{H}_{\delta}^{1}(A) &\leq \sum_{j=1,k=-\infty}^{\infty} \frac{2 \times diam\,(C_j \cap I_k)}{2} \\
&= \sum_{j=1}^{\infty} \left(\sum_{k=-\infty}^{\infty} diam\,(C_j \cap I_k) \right) \\
&= \sum_{j=1}^{\infty} diam\, C_j \qquad (4.0.4)
\end{aligned}$$

where the last step is from (4.0.3). Recall that (4.0.4) holds for any cover $\{C_j\}_{j=1}^{\infty}$ of A , then taking the infimum in (4.0.4) over these covers, we get

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

Thus, $\mathcal{H}_{\delta}^{1}(A) = \mathcal{L}^1(A)$ for any $\delta > 0$. Taking the limit as δ goes to 0, we get that $\mathcal{H}^1(A) = \mathcal{L}^1(A)$.

3. We fix an integer $m \geq 1$, and decompose the unit cube $Q \subseteq \mathbb{R}^n$ into m^n cubes with sides $\frac{1}{m}$ and diameter $\frac{\sqrt{n}}{m}$. Let $\delta = \frac{\sqrt{n}}{m}$; then

$$\begin{aligned}
\mathcal{H}_{\frac{\sqrt{n}}{m}}^{s}(Q) &\leq \sum_{i=1}^{m^n} \alpha(s) \left(\frac{\sqrt{n}}{2m} \right)^s \\
&\leq \sum_{i=1}^{m^n} \alpha(s) \left(\frac{\sqrt{n}}{m} \right)^s \\
&= \alpha(s) \sum_{i=1}^{m^n} \frac{n^{\frac{s}{2}}}{m^s} \\
&= \alpha(s) \sum_{i=1}^{m^n} m^{-s} n^{\frac{s}{2}} \\
&= \alpha(s)\, n^{\frac{s}{2}} m^{n-s}.
\end{aligned}$$

Let $m \to \infty$, we get $\mathcal{H}^s(Q) = 0$ and so by countable additivity $\mathcal{H}^s(\mathbb{R}^n) = 0$ for $s > n$.

4. Select sets C_j's such that $A \subseteq \bigcup_{j=1}^{\infty} C_j$. For any $\lambda > 0$, notice that

$$A \subseteq \bigcup_{j=1}^{\infty} C_j \Leftrightarrow \lambda A \subseteq \bigcup_{j=1}^{\infty} \lambda C_j.$$

Thus, there exists a 1-to-1 correspondance between covers of A and λA. Hence,

$$\begin{aligned}\mathcal{H}_\delta^s(\lambda A) &= \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,(\lambda C_j)}{2}\right)^s;\ \lambda A\subseteq\bigcup_{j=1}^{\infty}\lambda C_j\right\}\\ &= \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\lambda^s\left(\frac{diam\,C_j}{2}\right)^s;\ A\subseteq\bigcup_{j=1}^{\infty}C_j\right\}\\ &= \lambda^s\inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C_j}{2}\right)^s;\ A\subseteq\bigcup_{j=1}^{\infty}C_j\right\}\\ &= \lambda^s\mathcal{H}_\delta^s(A).\end{aligned}$$

5. Select sets C_j's such that $A\subseteq\bigcup_{j=1}^{\infty}C_j$. Notice that for any affine isometry $L:\mathbb{R}^n\longrightarrow\mathbb{R}^n$,

$$A\subseteq\bigcup_{j=1}^{\infty}C_j\Leftrightarrow L(A)\subseteq L\left(\bigcup_{j=1}^{\infty}C_j\right)=\bigcup_{j=1}^{\infty}L(C_j).$$

Thus, there exists a 1-to-1 correspondance between covers of A and $L(A)$, and hence

$$\begin{aligned}\mathcal{H}_\delta^s(L(A)) &= \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,(L(C_j))}{2}\right)^s;\ L(A)\subseteq\bigcup_{j=1}^{\infty}L(C_j)\right\}\\ &= \inf\left\{\sum_{j=1}^{\infty}\alpha(s)\left(\frac{diam\,C_j}{2}\right)^s;\ A\subseteq\bigcup_{j=1}^{\infty}C_j\right\}\\ &= \mathcal{H}_\delta^s(A).\end{aligned}$$

□

4.1 Isodiametric inequality

Throughout this section ,we will be proving that $\mathcal{H}^n=\mathcal{L}^n$ on $\mathbb{R}^n$. This cannot be seen easily since by definition, the Lebesgue measure $\mathcal{L}^n(A)$ is computed using arbitrary coverings of A , whereas the Hausdorff measure $\mathcal{H}^n(A)$ is computed in terms of arbitrary coverings of small diameter.

Lemma 4.1.1. *Let $f:\mathbb{R}^n\longrightarrow[0,\infty]$ be $\mathcal{L}^n$ measurable.Then the set*

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

which represents the region under the graph of f , is $\mathcal{L}^{n+1}$ measurable.

Proof. Let $B = \{x \in \mathbb{R}^n; f(x) = \infty\}$ and $C = \{x \in \mathbb{R}^n; 0 \leq f(x) < \infty\}$. In addition we define

$$\begin{aligned} C_{jk} &= \left\{x \in C;\ \frac{j}{k} \leq f(x) < \frac{j+1}{k},\ j \in \mathbb{N}, k \in \mathbb{N}^*\right\} \\ D_k &= \bigcup_{j=0}^{\infty} \left(C_{jk} \times [0, \frac{j}{k}]\right) \cup (B \times [0,\infty]) \quad , j \in \mathbb{N}, k \in \mathbb{N}^* \end{aligned}$$

and

$$E_k = \bigcup_{j=0}^{\infty} \left(C_{jk} \times [0, \frac{j+1}{k}]\right) \cup (B \times [0,\infty]) \quad , j \in \mathbb{N}, k \in \mathbb{N}^*$$

Since C_{jk} and B are $\mathcal{L}^n$ measurable in $\mathbb{R}^n$, and since $[0, \frac{j+1}{k}]$ and $[0,\infty]$ are $\mathcal{L}^1$ measurable, then E_k and D_k are $\mathcal{L}^{n+1}$ measurable. Moreover, $D_k \subset A \subset E_k$. Let us define $D = \bigcup_{k=1}^{\infty} D_k$ and $E = \bigcap_{k=1}^{\infty} E_k$. Then $D \subset A \subset E$ with D and E both $\mathcal{L}^{n+1}$ measurable.Now, since $D_k \subset D$ and $E \subseteq E_k$,then

$$E \setminus D = E \cap D^c \subseteq E_k \cap D_k^c = E_k \setminus D_k.$$

Denoting $\mathcal{B}^{n+1}(0,R) = B^n(0,R) \times [0,\infty]$, we get

$$\begin{aligned} \mathcal{L}^{n+1}\left((E \setminus D) \cap \mathcal{B}^{n+1}(0,R)\right) &\leq \mathcal{L}^{n+1}\left((E_k \setminus D_k) \cap \mathcal{B}^{n+1}(0,R)\right) \\ &= \mathcal{L}^{n+1}\left(\bigcup_{j=0}^{\infty} C_{jk} \times [\frac{j}{k}, \frac{j+1}{k}] \cap \mathcal{B}^{n+1}(0,R)\right) \\ &= \mathcal{L}^n\left(\left(\bigcup_{j=0}^{\infty} C_{jk}\right) \cap B^n(0,R)\right) \times \mathcal{L}^1\left([\frac{j}{k}, \frac{j+1}{k}]\right) \\ &\leq \frac{1}{k}\mathcal{L}^n\left(B^n(0,R)\right). \end{aligned}$$

Now as $K \to \infty$, the last term goes to zero and hence $\mathcal{L}^{n+1}\left((E \setminus D) \cap \mathcal{B}^{n+1}(0,R)\right) = 0$, which implies that

$$\begin{aligned} \mathcal{L}^{n+1}((E \setminus D)) &= \mathcal{L}^{n+1}\left((E \setminus D) \cap \mathbb{R}^{n+1}\right) \\ &= \mathcal{L}^{n+1}\left((E \setminus D) \cap \left(\bigcup_{n=1}^{\infty} \mathcal{B}^{n+1}(0,n)\right)\right) \\ &= \mathcal{L}^{n+1}\left(\bigcup_{n=1}^{\infty} \left((E \setminus D) \cap \mathcal{B}^{n+1}(0,n)\right)\right) \\ &\leq \sum_{n=1}^{\infty} \mathcal{L}^{n+1}\left((E \setminus D) \cap \mathcal{B}^{n+1}(0,n)\right) \\ &= 0. \end{aligned}$$

Hence, $\mathcal{L}^{n+1}((A \setminus D)) = 0$. $A \setminus D$ is $\mathcal{L}^{n+1}$ measurable (See Remark on page 2 in [1].) Since as noted earlier D is $\mathcal{L}^{n+1}$ measurable, then $A = (A \setminus D) \cup D$ is $\mathcal{L}^{n+1}$ measurable. $\square$

Notation *Fix* $a, b \in \mathbb{R}^n$, $|a| = 1$. *Let us define*

$$\begin{aligned} L_b^a &= \{b + ta;\ t \in \mathbb{R}\}\ ,\ \textit{the line through } b \textit{ in the direction of } a \\ \textit{and } P_a &= \{x \in \mathbb{R}^n;\ x.a = 0\}\ ,\ \textit{the plane through the origin perpendicular to } a. \end{aligned}$$

Definition 4.1.2. Let $a \in \mathbb{R}^n$, such that $|a| = 1$, and let $A \subset \mathbb{R}^n$. We define the Steiner Symmetrization of A with respect to the plane P_a to be the set

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

In the next lemma, we prove some properties of Steiner Symmetrization.

Lemma 4.1.3. *Let* $A \subset \mathbb{R}^n$ *be a closed set.*

1. $diam\, S_a(A) \leq diam\, A$.

2. $S_a(A)$ *is* $\mathcal{L}^n$ *measurable and* $\mathcal{L}^n(S_a(A)) = \mathcal{L}^n(A)$.

Proof. 1. If $diam\ A = \infty$ then statement (1) is trivial, so we will assume that $diam\ A < \infty$. Fix $\epsilon > 0$ and by definition of supremum, select $x, y \in S_a(A)$ such that

$$diam\ S_a(A) \leq |x - y| + \epsilon. \tag{4.1.1}$$

Let

$$b = x - (x.a).a \textit{ and } c = y - (y.a).a.$$

Moreover, $b \in P_a$ since

$$\begin{aligned} b.a &= (x - (x.a).a).a \\ &= \left(x - x.|a|^2\right).a \\ &= (x - x).a \\ &= 0. \end{aligned} \tag{4.1.2}$$

Similarly, $c \in P_a$.
Note that by definition of S_a , we have

$$|x.a| > \frac{1}{2}\mathcal{H}^1\left(A \cap L_b^a\right) \tag{4.1.3}$$

and

$$|y.a| \;>\; \frac{1}{2}\mathcal{H}^1\left(A \cap L_b^a\right). \tag{4.1.4}$$

Now set

$$\begin{aligned} r &= \inf\{t;\ b + ta \in A\} \\ s &= \sup\{t;\ b + ta \in A\} \\ u &= \inf\{t;\ c + ta \in A\} \\ v &= \sup\{t;\ c + ta \in A\} \end{aligned}$$

Assume that $v - r \geq s - u$. Then $\frac{1}{2}(v - r) \geq \frac{1}{2}(s - u)$. Using the fact that $\frac{1}{2}(v - r) = (v - r) - \frac{1}{2}(v - r)$; we get

$$(v - r) - \frac{1}{2}(v - r) \;\geq\; \frac{1}{2}(s - u). \tag{4.1.5}$$

Adding $\frac{1}{2}(v - r)$ on both sides of (4.1.5) we get

$$\begin{aligned} (v - r) &\geq \frac{1}{2}(v - r) + \frac{1}{2}(s - u) \\ &= \frac{1}{2}(s - r) + \frac{1}{2}(v - u). \end{aligned} \tag{4.1.6}$$

However,

$$\begin{aligned} s - r &= \sup\{t;\ b + ta \in A\} - \inf\{t;\ b + ta \in A\} \geq \mathcal{H}^1\left(A \cap L_b^a\right) \\ &\textit{and} \\ v - u &= \sup\{t;\ c + ta \in A\} - \inf\{t;\ c + ta \in A\} \geq \mathcal{H}^1\left(A \cap L_c^a\right). \end{aligned}$$

Thus, plugging in (4.1.6), we get

$$\begin{aligned} v - r &\geq \frac{1}{2}\mathcal{H}^1\left(A \cap L_b^a\right) + \frac{1}{2}\mathcal{H}^1\left(A \cap L_c^a\right) \\ &\geq |x.a| + |y.a| \\ &\geq |x.a - y.a|, \end{aligned} \tag{4.1.7}$$

where we used (4.1.3) and (4.1.4) in the step before the last. Now, recall by (4.1.1) that

$$\mathit{diam}\, S_a - \epsilon \leq |x - y|, \tag{4.1.8}$$

and by definition of b and c that $x = b + (x.a)\,.a$ and $y = c + (y.a)\,.a$. This means that

$$x - y = b - c + ((x.a) - (y.a)). \tag{4.1.9}$$

Notice that $(b-c)$ is perpendicular to $(x-y)$ since their dot product gives 0. Hence, by squaring both sides of the equation (4.1.9) and using pythagorian rule we get

$$|x-y|^2 = |b-c|^2 + |((x.a)-(y.a)).a|^2. \tag{4.1.10}$$

Therefore, combining (4.1.8) , (4.1.10) and (4.1.7), we get

$$\begin{aligned} (diam\ S_a - \epsilon)^2 &\leq |x-y|^2 \\ &= |b-c|^2 + |((x.a)-(y.a))|^2 \\ &\leq |b-c|^2 + |(v-r)|^2 \end{aligned} \tag{4.1.11}$$

However note that,using (4.1.2) we have

$$\begin{aligned} |(b+ra)-(c+va)|^2 &= |b+ra-c-va|^2 \\ &= |(b-c)+(r-v).a|^2 \\ &= |b-c|^2 + |(r-v).a|^2 + 2(b-c)(r-v).a \\ &= |b-c|^2 + |r-v|^2 + 2b(r-v).a - 2c(r-v).a \\ &= |b-c|^2 + |(v-r)|^2. \end{aligned} \tag{4.1.12}$$

So plugging (4.1.12) in (4.1.11) we get

$$(diam\ S_a - \epsilon)^2 \leq |(b+ra)-(c+va)|^2.$$

Since A is closed , and $v = \sup\{t;\ c+ta \in A\}$, then $c+va \in A$. Similarly, $b+ra \in A$. Thus (4.1.11) becomes

$$(diam\ S_a - \epsilon)^2 \leq (diam\ A)^2$$

and hence,

$$diam\ S_a - \epsilon \leq diam\, A.$$

2. Let $A \subset \mathbb{R}^n$ be a closed set.We start by studying $\mathcal{L}^n(A)$. Since $\mathcal{L}^n$ is rotation invariant , then without loss of generality take $a = e_n = (0,0,\cdots,1)$, making $P_a = P_{e_n} = \mathbb{R}^{n-1}$. Knowing that $\mathcal{H}^1 = \mathcal{L}^1$ on $\mathbb{R}^1$ and $\mathcal{L}^n = \mathcal{L}^1 \times \mathcal{L}^{n-1}$ then we get

$$\begin{aligned} \mathcal{L}^n(A) &= \int \chi_A\, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^{n-1}\times\mathbb{R}} \chi_A(x,y)\ d\mathcal{L}^n(x,y) \\ &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \chi_A(x,y)\ d\mathcal{L}^1(y)\ d\mathcal{L}^{n-1}(x), \end{aligned} \tag{4.1.13}$$

where in the last step we used Fubini's theorem (see theorem 3.0.5) . Notice that $\chi_A(x,y) = \begin{cases} 1; & (x,y) \in A \\ 0; & (x,y) \notin A. \end{cases}$

Now let $A_x = \{y \in \mathbb{R};\ (x,y) \in A\}$. Then $\chi_{A_x}(y) = \begin{cases} 1; & y \in A_x \\ 0; & y \notin A_x \end{cases} = \begin{cases} 1; & (x,y) \in A \\ 0; & (x,y) \notin A \end{cases} =$
$\chi_A(x,y)$.
Since the inner integral of (4.1.13) is independent of x ,(4.1.13) becomes

$$\begin{aligned} \mathcal{L}^n(A) &= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{A_x}(y)\ d\mathcal{L}^1(y) \right) d\mathcal{L}^{n-1}(x) \\ &= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1(A_x)\ d\mathcal{L}^{n-1}(x). \end{aligned}$$

Let the map $f : \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$ defined by $f(b) = \mathcal{H}^1(A \cap L_b^a)$ be $\mathcal{L}^{n-1}$ measurable. Since $\mathcal{L}^1$ is translation invariant, then

$$\begin{aligned} \mathcal{L}^n(A) &= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1(A \cap L_b^a)\ d\mathcal{L}^{n-1}(b) \\ &= \int_{\mathbb{R}^{n-1}} \mathcal{H}^1(A \cap L_b^a)\ d\mathcal{L}^{n-1}(b) \\ &= \int_{\mathbb{R}^{n-1}} f(b)\ db. \end{aligned} \tag{4.1.14}$$

On the other hand,

$$\begin{aligned} S_a(A) &= \bigcup_{b \in \mathbb{R}^{n-1}, A \cap L_b^a \neq \phi} \left\{ b + ta;\ |t| < \frac{1}{2}\mathcal{H}^1(A \cap L_b^a) \right\} \\ &= \left\{ (b,y);\ -\frac{1}{2}\mathcal{H}^1(A \cap L_b^a) \le y \le \frac{1}{2}\mathcal{H}^1(A \cap L_b^a) \right\} \setminus \left\{ (b,0);\ A \cap L_b^a = \phi \right\} \\ &= \left\{ (b,y);\ -\frac{f(b)}{2} \le y \le \frac{f(b)}{2} \right\} \setminus \left\{ (b,0);\ A \cap L_b^a = \phi \right\} \\ &= \left\{ (b,y);\ b \in \mathbb{R}^{n-1}, y \in \mathbb{R}, 0 \le y \le \frac{f(b)}{2} \right\} \bigcup \left\{ (b,y);\ b \in \mathbb{R}^{n-1}, y \in \mathbb{R}, -\frac{f(b)}{2} \le y \le 0 \right\} \\ &\qquad \setminus \left\{ (b,0);\ A \cap L_b^a = \phi \right\}. \end{aligned}$$

Using lemma (4.1.1), we get that the first part of the union is $\mathcal{L}^n$ measurable. But the second part of the union is nothing but the reflection with respect to $\mathbb{R}^{n-1}$ of the first part, and hence is $\mathcal{L}^n$ measurable also. Moreover, to see that $B := \left\{ (b,0);\ A \cap L_b^a = \phi \right\}$ is measurable; notice that

$$\left\{ (b,0);\ A \cap L_b^a = \phi \right\} = B \subseteq \mathbb{R}^{n-1}.$$

That is

$$\left\{ (b,0);\ A \cap L_b^a \neq \phi \right\} = B^c = pr_{\mathbb{R}^{n-1}}(A).$$

Hence, $B^c = \mathbb{R}^{n-1} \setminus B$, which is measurable. Thus, $S_a(A)$ is $\mathcal{L}^n$ measurable.To see this , let $B = \{b \in \mathbb{R}^{n-1};\ A \cap L_b^a \neq \phi\}$. Then $B^c = \{b \in \mathbb{R}^{n-1};\ A \cap L_b^a = \phi\}$. So if $b \in B$

then $f(b)$ is $f(b)$ and if $b \in B^c$ then $f(b) = 0$. Hence,

$$\begin{aligned}
\mathcal{L}^n(S_a(A)) &= \mathcal{L}^n\left(\left\{ (b,y);\ b \in B\ , y \in \mathbb{R}, -\frac{f(b)}{2} \le y \le \frac{f(b)}{2} \right\}\right) \\
&= \mathcal{L}^n\left(B \times \left[-\frac{f(b)}{2}, \frac{f(b)}{2}\right]\right) \\
&= \int_{\mathbb{R}^n} \chi_{B\times[-\frac{f(b)}{2},\frac{f(b)}{2}]}(b,y) \\
&= \int_{\mathbb{R}^{n-1}\times\mathbb{R}} \chi_B(b) \cdot \chi_{[-\frac{f(b)}{2},\frac{f(b)}{2}]}(y). \qquad (4.1.15)
\end{aligned}$$

Using Fubini's theorem (see theorem 3.0.5), we get

$$\begin{aligned}
\mathcal{L}^n(S_a(A)) &= \int_{\mathbb{R}^{n-1}} \chi_B(b) \int_{\mathbb{R}} \left(\chi_{[-\frac{f(b)}{2},\frac{f(b)}{2}]}(y)\ d\mathcal{L}^1(y)\right)\ d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} \chi_B(b) f(b)\ d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} (\chi_B(b) f(b) + \chi_{B^c}(b) f(b))\, d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} (\chi_B + \chi_{B^c}) f(b)\, d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} \chi_{\mathbb{R}^{n-1}} f(b)\ d\mathcal{L}^{n-1}(b) \\
&= \int_{\mathbb{R}^{n-1}} f(b)\ d\mathcal{L}^{n-1}(b) \\
&= \mathcal{L}^n(A).
\end{aligned}$$

□

Theorem 4.1.4. *Isodiametric Inequality*
For all sets $A \subset \mathbb{R}^n$, $\mathcal{L}^n(A) \le \alpha(n)\left(\frac{diam A}{2}\right)^n$.

Proof. If $diam\, A = \infty$ then it is trivial. Let us assume that $diam\, A < \infty$. Let $\{e_1, \cdots, e_n\}$ be the standard basis for $\mathbb{R}^n$, define $A_1 = S_{e_1}(A)$, $A_2 = S_{e_2}(A_1)$, $\cdots$, $A_n = S_{e_n}(A_{n-1})$, and write $A^* = A_n$.

1. **Claim # 1**: A^* is symmetric with respect to the origin.
Proof of Claim # 1: We show this by induction. By definition of the Steiner symmetrization, A_1 is symmetric with respect to P_{e_1}. Now let $1 \le k < n$ and assume that A_k is symmetric with respect to $P_{e_1}, P_{e_2}, \cdots, P_{e_k}$. We prove A_{k+1} is symmetric with respect to $P_{e_1}, P_{e_2}, \cdots, P_{e_{k+1}}$. By definition, $A_{k+1} = S_{e_{k+1}}(A_k)$ is symmetric with respect to $P_{e_{k+1}}$. Fix $1 \le j \le k$ and let $S_j : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ to be the reflection through P_{e_j}. Fix $b \in P_{e_{k+1}}$. Since A_k is symmetric with respect to P_{e_j}, then $S_j(A_k) = A_k$ and

we get $\mathcal{H}^1\left(A_k \cap L_b^{e_k+1}\right) = \mathcal{H}^1\left(A_k \cap L_{S_j b}^{e_k+1}\right)$. Since $A_{k+1} = S_{e_{k+1}}\left(A_k\right)$, by definition of S_{e_k}, the latter equality implies that $\left\{t;\ b + te_{k+1} \in A_{k+1}\right\} = \left\{t;\ S_j b + te_{k+1} \in A_{k+1}\right\}$ that is, $S_j\left(A_{k+1}\right) = A_{k+1}$. So, A_{k+1} is symmetric with respect to P_{e_j}, thus $A^* = A_n$ is symmetric with respect to $P_{e_1}, P_{e_2}, \cdots, P_{e_n}$ and hence with respect to the origin.

2. **Claim # 2** $\mathcal{L}^n\left(A^*\right) \leq \alpha(n)\left(\dfrac{diam\, A^*}{2}\right)^n$.

   Proof of Claim # 2: Let $x \in A^*$ then by **Claim # 1**, we get that $-x \in A^*$. Thus, $diam\, A^* \geq 2|x|$ that is $|x| \leq \dfrac{diam\, A^*}{2}$. This implies that $A^* \subset B\left(0, \frac{diam\, A^*}{2}\right)$. Therefore,

$$\mathcal{L}^n\left(A^*\right) \leq \mathcal{L}^n\left(B\left(0, \frac{diam\ A^*}{2}\right)\right) = \alpha(n)\left(\frac{diam\, A^*}{2}\right)^n. \qquad (4.1.16)$$

3. **Claim # 3** $\mathcal{L}^n(A) \leq \alpha(n)\left(\dfrac{diam\, A}{2}\right)^n$.

   Proof of claim # 3:$\overline{A}$ is $\mathcal{L}^n$ measurable since it is closed, and thus by applying lemma 4.1.3 n times on $\bar{A}$ we get,

$$\mathcal{L}^n\left(\overline{A}\right) = \mathcal{L}^n\left(S_{e_1}\left(\overline{A}\right)\right) = \mathcal{L}^n\left(\overline{A}_1\right) = \mathcal{L}^n\left(S_{e_2}\left(\overline{A}_1\right)\right) = \mathcal{L}^n\left(\overline{A}_2\right) = \cdots = \mathcal{L}^n\left(\overline{A}_n\right) = \mathcal{L}^n\left(\overline{A}\right)^*.$$

   Moreover by applying lemma 4.1.3 n times, we get

$$diam\ \left(\bar{A}\right)^* \leq diam\left(\bar{A}\right) \qquad (4.1.17)$$

   , and hence

$$\begin{aligned} \mathcal{L}^n(A) &\leq \mathcal{L}^n\left(\bar{A}\right) \\ &= \mathcal{L}^n\left(\left(\bar{A}\right)^*\right) \\ &\leq \alpha(n)\left(\frac{diam\ \left(\bar{A}\right)^*}{2}\right)^n \end{aligned} \qquad (4.1.18)$$

   where the last equality comes from **Claim #2** used on $\bar{A}$. Using (4.1.17) on (4.1.18) we get

$$\begin{aligned} \mathcal{L}^n(A) &\leq \alpha(n)\left(\frac{diam\ \bar{A}}{2}\right)^n \\ &= \alpha(n)\left(\frac{diam\ A}{2}\right)^n. \end{aligned}$$

□

Theorem 4.1.5. *$\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$*

Proof. Let $A \subset \mathbb{R}^n$.
Claim # 1: $\mathcal{L}^n(A) \le \mathcal{H}^n(A)$.
Proof of claim # 1:Fix $\delta > 0$. Choose sets $\{C_j\}_{j=1}^{\infty}$ that cover A and such that $diam\ C_j \le \delta$. By countable subadditivity we get

$$\mathcal{L}^n(A) \le \sum_{j=1}^{\infty} \mathcal{L}^n(C_j)$$

and hence using the isodiametric inequality we get

$$\mathcal{L}^n(A) \le \sum_{j=1}^{\infty} \alpha(n) \left(\frac{diam\ C_j}{2}\right)^n.$$

Taking the infimum over all such sets $\{C_j\}_{j=1}^{\infty}$ we get

$$\begin{aligned} \mathcal{L}^n(A) &\le \inf \sum_{j=1}^{\infty} \alpha(n) \left(\frac{diam\ C_j}{2}\right)^n \\ &= \mathcal{H}^n_\delta(A) \end{aligned}$$

Thus,

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

Before moving to **Claim # 2**, recall that $\mathcal{L}^n$ is the product of $\mathcal{L}^1 \times \mathcal{L}^1 \times \cdots \times \mathcal{L}^1$ (n times). (See theorem 3.0.6). Moreover we know by the definition of Lebesgue measure that for all $A \subset \mathbb{R}^n$ and $\delta > 0$

$$\mathcal{L}^n(A) = \inf\left\{ \sum_{i=1}^{\infty} \mathcal{L}^n(Q_i)\,; Q_i \text{ are cubes }; A \subset \bigcup_{i=1}^{\infty} Q_i\ , diam\, Q_i \le \delta \right\}.$$

Claim # 2:$\mathcal{H}^n$ is absolutely continuous with respect to $\mathcal{L}^n$.
Proof of claim # 2: For each cube $Q \subset \mathbb{R}^n$ of side s we have,

$$\begin{aligned} \mathcal{L}^n(Q) &= s^n \\ &= \left(\frac{\sqrt{n}s}{\sqrt{n}}\right)^n \\ &= \left(\frac{diam\, Q}{\sqrt{n}}\right)^n. \end{aligned}$$

Let $C_n = \alpha(n) \left(\frac{\sqrt{n}}{2}\right)^n$. Then,

$$C_n \mathcal{L}^n(Q) = \alpha(n) \left(\frac{diam\, Q}{2}\right)^n. \tag{4.1.19}$$

Moreover, notice that the set of all covers of $A \subset \mathbb{R}^n$ by cubes Q_j of $diam\ Q_j < \delta$ is subset to the set of all covers C_j of A such that $diam\, C_j \leq \delta$. Hence,

$$\begin{aligned} \mathcal{H}^n_\delta (A) &\leq \inf \left\{ \sum_{j=1}^{\infty} \alpha(n) \left(\frac{diam\ Q_i}{2} \right)^n ; \ Q_i \text{ cubes }; A \subset \bigcup_{i=1}^{\infty} Q_i \,, diam\, Q_i \leq \delta \right\} \\ &= C_n \inf \left\{ \sum_{j=1}^{\infty} \mathcal{L}^n (Q_j) ; \ Q_i \text{ are cubes }; A \subset \bigcup_{i=1}^{\infty} Q_i \,, diam\, Q_i \leq \delta \right\} \\ &= C_n \mathcal{L}^n (A) \end{aligned}$$

where we used (4.1.19) in the step before the last. Let $\delta \to 0$, we get

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

Claim #3:$\mathcal{H}^n (A) \leq \mathcal{L}^n (A)$ for all $A \subset \mathbb{R}^n$.

Proof of claim # 3: Fix $\delta, \epsilon > 0$. By the definition of infimum, choose cubes $\{Q_i\}_{i=1}^{\infty}$ such that $A \subset \bigcup_{i=1}^{\infty} Q_i$, $diam\, Q_i < \delta$ and

$$\sum_{i=1}^{\infty} \mathcal{L}^n (Q_i) \leq \mathcal{L}^n (A) + \epsilon. \tag{4.1.20}$$

Using Vitali's covering (see theorem 3.0.7) we get that for each $i \in \mathbb{N}$ there exist disjoint closed balls $\{B^i_k\}_{k=1}^{\infty}$ contained in Q^0_i , the interior of Q_i , such that $diam\, B^i_k \leq \delta$ and

$$\mathcal{L}^n \left(Q_i \setminus \bigcup_{k=1}^{\infty} B^i_k \right) = \mathcal{L}^n \left(Q^0_i \setminus \bigcup_{k=1}^{\infty} B^i_k \right) = 0.$$

Using **Claim #2** we get $\mathcal{H}^n \left(Q_i \setminus \bigcup_{k=1}^{\infty} B^i_k \right) = 0$. Also $Q_i = \left(Q_i \setminus \bigcup_{k=1}^{\infty} B^i_k \right) \cup \left(\bigcup_{k=1}^{\infty} B^i_k \right)$. Thus by countable additivity , we get

$$\mathcal{H}^n (Q_i) = \mathcal{H}^n \left(Q_i \setminus \bigcup_{k=1}^{\infty} B^i_k \right) + \mathcal{H}^n \left(\bigcup_{k=1}^{\infty} B^i_k \right)$$

which gives us that

$$\mathcal{H}^n (Q_i) = \mathcal{H}^n \left(\bigcup_{k=1}^{\infty} B^i_k \right) . \tag{4.1.21}$$

Now, by (4.1.21) and countable subadditivity, we have

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

Where we used (4.1.20) in the last step. Let $\epsilon \to 0$, we get $\mathcal{H}^n_\delta (A) \leq \mathcal{L}^n (A)$. Let $\delta \to 0$, we get $\mathcal{H}^n (A) \leq \mathcal{L}^n (A)$ and we are done. □

4.2 Hausdorff measure and Lipschitz mappings

Definition 4.2.1. A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is called a Lipschitz function if there exists a constant C such that $|f(x) - f(y)| \leq C|x - y|$ for all x and y in $\mathbb{R}^n$.

Definition 4.2.2. Let f be a Lipschitz function. Define $Lip(f) = \sup \left\{ \frac{|f(x) - f(y)|}{|x - y|} ; x, y \in \mathbb{R}^n , x \neq y \right\}$. We call $Lip(f)$ the Lipschitz constant of the function f.

Theorem 4.2.3. *Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a Lipschitz function , $A \subset \mathbb{R}^n$, $0 \leq s < \infty$. Then,*

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

Proof. Fix $\delta > 0$. Choose sets $\{C_i\}_{i=1}^{\infty} \subset \mathbb{R}^n$ such that $diam\, C_i \leq \delta$ and $A \subset \bigcup_{i=1}^{\infty} C_i$. Let $x, y \in C_i$ then,

$$\begin{aligned}
diam\, C_i &= \sup \left\{ |x - y|,\ x, y \in C_i \right\} \\
&\textit{and} \\
diam\, f(C_i) &= \sup \left\{ |f(x) - f(y)|,\ x, y \in C_i \right\}.
\end{aligned}$$

Notice that

$$\begin{aligned} |f(x) - f(y)| &= \frac{|f(x) - f(y)|}{|x-y|}|x-y| \\ &\leq \frac{|f(x) - f(y)|}{|x-y|}\sup|x-y|. \end{aligned}$$

Taking the supremum on both sides, we get

$$\sup|f(x) - f(y)| \leq \sup\left(\frac{|f(x) - f(y)|}{|x-y|}\right)\sup(|x-y|).$$

This shows that

$$diam\, f(C_i) \leq Lip(f)\ diam\, C_i \leq Lip(f)\ \delta.$$

Now,

$$f(A) \subseteq f\left(\bigcup_{i=1}^{\infty} C_i\right) \subseteq \bigcup_{i=1}^{\infty} f(C_i).$$

Thus, $\{f(C_i)\}_{i=1}^{\infty}$ are a cover for $f(A)$ with $diam\, f(C_i) \leq Lip\, f\ \delta$. Hence,

$$\begin{aligned} \mathcal{H}^s_{Lip(f)\delta}(f(A)) &\leq \sum_{i=1}^{\infty} \alpha(s)\left(\frac{diam\, f(C_i)}{2}\right)^s \\ &\leq (Lip(f))^s \sum_{i=1}^{\infty} \alpha(s)\left(\frac{diam\, C_i}{2}\right)^s \end{aligned}$$

Taking the infimum over all such sets C_i we get

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

Letting $\delta \to 0$ we get

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

□

Corollary 4.2.4. *Suppose $n > k$. Let $P : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be the usual projection. Let $A \subset \mathbb{R}^n$, $0 \leq s < \infty$, then $\mathcal{H}^s(P(A)) \leq \mathcal{H}^s(A)$.*

Proof. The projection function is a Lipschitz function with $Lip(P) = 1$. To see that, take $x, y \in \mathbb{R}^n$. Since the projection function is linear, with norm 1 then, $|P(x) - P(y)| = |P(x-y)| \leq |x-y|$. Hence $\frac{|P(x) - P(y)|}{|x-y|} \leq 1$ which implies that $Lip(P) \leq 1$. To see that $Lip(P) = 1$ take $x \in \mathbb{R}^n = (x_1, \cdots, x_k, 0, \cdots, 0)$ and $y \in \mathbb{R}^n = (y_1, \cdots, y_k, 0, \cdots, 0)$ then $P(x) = x$ and $P(y) = y$. Thus $|P(x) - P(y)| = |x-y|$ that is, $\frac{|P(x) - P(y)|}{|x-y|} = 1$. Using theorem 4.1.5 we get that

$$\begin{aligned} \mathcal{H}^s(P(A)) &\leq Lip(P)\, \mathcal{H}^s(A) \\ &= \mathcal{H}^s(A). \end{aligned}$$

□

Chapter 5

Lipschitz functions , Rademacher's Theorem

Rademacher's Theorem states that Lipschitz functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ are differentiable $\mathcal{L}^n$ *a.e.* To be able to state Rademacher's Theorem, we need to define what it means for a function to be Lipschitz and to define differentiability from $\mathbb{R}^n$ to $\mathbb{R}^m$.
We start by defining Lipschitz functions and locally Lipschitz functions.

Definition 5.0.1. 1. A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(n \leq m)$ is said to be Lipschitz if

$$|f(x) - f(y)| \leq C|x - y| \tag{5.0.1}$$

for some constant C and for all x and y in $\mathbb{R}^n$. Define

$$Lip(f) = \sup\left\{\frac{|f(x) - f(y)|}{|x - y|}; \forall x, y \in \mathbb{R}^n, x \neq y\right\};$$

Note that $Lip(f)$ is the smallest constant C such that (5.0.1) holds for all x and y.

2. A function $f : A \longrightarrow \mathbb{R}^m$ $(A \subset \mathbb{R}^n)$ is said to be locally Lipschitz if, for each compact set $K \subset A$,there exists a constant C_k such that

$$|f(x) - f(y)| \leq C_k|x - y| \ \forall x, y \in K.$$

Theorem 5.0.2. *Extension of Lipschitz functions*
Suppose $f : A \longrightarrow \mathbb{R}^m$ is a Lipschitz function where $A \subset \mathbb{R}^n$, then there exists a Lipschitz function $\bar{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that :

1. $\bar{f} = f$ on A.

2. $Lip(\bar{f}) \leq \sqrt{m}\, Lip(f)$.

Proof. Let us prove the theorem first for the case if $f : A \longrightarrow \mathbb{R}$.
We start by showing that $\bar{f}$ is Lipschitz. Define $\bar{f}(x) = \inf_{a \in A} \left\{ f(a) + Lip(f)|x - a| \right\}$. Let $x, y \in \mathbb{R}^n$ then

$$\begin{aligned}\bar{f}(x) &\leq \inf_{a \in A} \left\{ f(a) + Lip\ (f)\left(|y - a| + |x - y|\right) \right\} \\ &= \bar{f}(y) + Lip\ (f)|x - y|\end{aligned}$$

and similarly ,

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

To show (1), let $b \in A$. Notice that $\bar{f}(b) = \inf_{a \in A} \left\{ f(a) + Lip(f)|b - a| \right\}$. However, $b \in A$ since $\bar{f}(b) \leq f(b) + Lip(f)|b - b| = f(b)$ and hence $\bar{f}(b) \leq f(b)$.
Conversely, for all a in A we have $|f(b) - f(a)| \leq Lip(f)|b - a|$. This implies that

$$-Lip(f)|b - a| \leq f(b) - f(a) \leq Lip(f)|b - a|.$$

Hence, $f(b) \leq f(a) + Lip(f)|b - a|$. But $\bar{f}(b) = \inf_{a \in A} \left\{ f(a) + Lip(f)|b - a| \right\}$. Thus, $f(b) \leq \bar{f}(b)$.
To show (2), Let $f : A \longrightarrow \mathbb{R}$, such that $A \subset \mathbb{R}^n$. Then for all $x, y \in \mathbb{R}^n$ we have $|\bar{f}(x) - \bar{f}(y)| \leq Lip(f)|x - y|$. This implies

$$\begin{aligned}\frac{|\bar{f}(x) - \bar{f}(y)|}{|x - y|} &\leq Lip(f) \\ &= \sqrt{m}\, Lip(f) \ \text{(since m is equal to 1 in this case.)}\end{aligned}$$

Hence,

$$\begin{aligned}Lip(\bar{f}) &= \sup_{x, y \in \mathbb{R}^n} \frac{|\bar{f}(x) - \bar{f}(y)|}{|x - y|} \\ &\leq \sqrt{m}\, Lip(f).\end{aligned}$$

For the general case, let $f : A \longrightarrow \mathbb{R}^m$.
$x \longrightarrow f(x) = (f_1(x), \cdots, f_m(x))$ be a Lipschitz function.
Notice that each $f_i : A \longrightarrow \mathbb{R}$
$x \longrightarrow f_i(x)$ is Lipschitz with $Lip(f_i) \leq Lip(f)$, since $|f_i(x) - f_i(y)| < |f(x) - f(y)| < Lip(f)|x - y|$ for all $x, y \in \mathbb{R}^n$.
Thus, by our discussion above, we can extend f_i to $\bar{f}_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\begin{cases} \bar{f}_i = f_i \text{ on } A \\ Lip(\bar{f}_i) \leq Lip(f_i) \end{cases}$

Then, we have

$$\begin{aligned} |\bar{f}(x) - \bar{f}(y)|^2 &= \sum_{i=1}^{m} |\bar{f}_i(x) - \bar{f}_i(y)|^2 \\ &\leq \sum_{i=1}^{m} (Lip\, f_i)^2 |x-y|^2 \\ &\leq \sum_{i=1}^{m} (Lip\, f)^2 |x-y|^2 \\ &= m\, (Lip\, f)^2 |x-y|^2 \end{aligned}$$

which implies $\dfrac{|\bar{f}(x) - \bar{f}(y)|^2}{|x-y|^2} \leq m\,(Lip\, f)^2$.

Thus, $\dfrac{\bar{f}(x) - \bar{f}(y)}{|x-y|} \leq \sqrt{m}\, Lip\,(f)$ for all $x, y \in \mathbb{R}^n$, which implies that $\sup\limits_{x,y \in \mathbb{R}^n} \dfrac{|\bar{f}(x) - \bar{f}(y)|}{|\,x-y\,|} \leq$ $\sqrt{m}\, Lip\,(f)$. Consequently, $Lip\,(\bar{f}) \leq \sqrt{m}\, Lip\,(f)$. □

Next we define differentiability for functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

Definition 5.0.3. A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $x \in \mathbb{R}^n$, if there exists a linear mapping $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{y\to x} \frac{|f(y) - f(x) - L(x-y)|}{|x-y|} = 0. \tag{5.0.2}$$

Remark 5.0.4. Let us prove that if such a linear map exists, it is unique and we write $Df(x)$ for L. We call $Df(x)$ is the derivative of f at x.

Proof. Suppose there exists 2 linear functions L_1 and L_2 such that $L_1, L_2 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, and that satisfy the equation above (5.0.2). Fix x and v such that $|v| = 1$. Let $y = x + tv$. Then, $|y - x| = |tv| = |t|$ then, $x - y = -tv$. Hence,

$$\lim_{t\to 0} \left| \frac{f(x+tv) - f(x) - L_1(-tv)}{|t|} \right| = 0$$

and

$$\lim_{t\to 0} \left| \frac{f(x+tv) - f(x) - L_2(-tv)}{|t|} \right| = 0.$$

This implies that $\lim_{t\to 0} \left| \frac{f(x+tv) - f(x)}{|t|} + \frac{L_1(tv)}{|t|} \right| = \lim_{t\to 0} \left| \frac{f(x+tv) - f(x)}{|t|} + \frac{L_2(tv)}{|t|} \right|$. Then, $\lim_{t\to 0} \left| \frac{L_1(tv)}{|t|} - \frac{L_2(tv)}{|t|} \right| = 0$ and consequently $\lim_{t\to 0} \left| \frac{t[L_1 v - L_2 v]}{|t|} \right| = 0$. This in return gives $\lim_{t\to 0} |L_1 v - L_2 v| = 0$. Thus, $|L_1 v - L_2 v| = 0$. Hence,

$$L_1(v) = L_2(v) \tag{5.0.3}$$

for all $v \in \mathbb{R}^n$ such that $|v| = 1$. In general, Let $x \in \mathbb{R}^n$, write $x = \frac{x}{|x|} . |x|$ then using linearity of L_1 and L_2 and (5.0.3) for $v = \frac{x}{|x|}$, we get

$$\begin{aligned} L_1(x) &= L_1\left(\frac{x}{|x|} . |x|\right) \\ &= |x| \, L_1\left(\frac{x}{|x|}\right) \\ &= |x| \, L_2\left(\frac{x}{|x|}\right) \\ &= L_2\left(\frac{x}{|x|} . |x|\right) \\ &= L_2(x). \end{aligned}$$

□

Theorem 5.0.5. *Rademacher's theorem*
Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a locally Lipschitz function, then f is differentiable $\mathcal{L}^n$ a.e.

Proof. ***Case 1***: Assume $m = 1$; since differentiability is a local property, we may assume that f is Lipschitz. Fix any $v \in \mathbb{R}^n$ such that $|v| = 1$. For $x \in \mathbb{R}^n$ define

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

if this limit exists. (This is the directional derivative of f at x with the direction of v).
Claim # 1 : $D_v f(x)$ exists for $\mathcal{L}^n$ *a.e.* x.
Proof of Claim # 1 : Since f is a continuous function then,

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

is Borel measurable and

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

is also Borel measurable. To see this, given that f is a lipschitz function then it is continuous. Fix $t > 0$, $v \in \mathbb{R}^n$,then $f(x+tv)$ is also continuous by translation and $\dfrac{f(x+tv)}{t}$ is also continuous by dilation. So $g(x) = \dfrac{f(x+tv) - f(x)}{t}$ is continuous which implies that g is borel; because $\left(g^{-1}(\textit{open set}) \text{ is an open set which is borel}\right)$. Hence $\overline{\lim}_{t \to 0}\, g(x)$ is borel and the same goes for $\underline{\lim}_{t \to 0}\, g(x)$. Thus we get that
$A_v := \{x \in \mathbb{R}^n \text{ such that } D_v f(x) \text{ does not exist}\}$
$= \{x \in \mathbb{R}^n \text{ such that } \underline{D}_v f(x) < \overline{D}_v f(x)\}$ is Borel measurable.

Now for each $x, v \in \mathbb{R}^n$, with $|v| = 1$, define $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ by $\varphi(t) = f(x+tv)$ where $t \in \mathbb{R}$. Let us show that φ is Lipschitz, and absolutely continuous, thus making it differentiable $\mathcal{L}^n$ *a.e.*

- φ **is Lispchitz:** $\varphi(a) - \varphi(b) = f(x+av) - f(x+bv)$; but f is Lipschitz then

$$\begin{aligned} |f(x+av) - f(x+bv)| &\leq C|x+av-x-bv| \\ &\leq C|av-bv| \\ &= C|v(a-b)| \\ &\leq C|v||a-b| \\ &= C|a-b|. \end{aligned}$$

- φ **is absolutely continuous:** Since φ is a Lipschitz function then it is absolutely continous (see Theorem 3.0.3). Since φ is absolutely continuous then φ' exists $\mathcal{L}^1$ a.e

, that is f is differentiable $\mathcal{L}^1$ a.e on any line L parrallel to v. Consequently,

$$\begin{aligned} A_v \cap L &= \Big\{x \ \in \mathbb{R}^n \text{ such that } f \text{ is not differentiable at } x\Big\} \cap L \\ &= \{x \ \in L \text{ such that } f \text{ is not differentiable at } x\}. \end{aligned}$$

Which implies that $\mathcal{L}^1 (A_v \cap L) = 0$; hence, $\mathcal{H}^1 (A_v \cap L) = 0$ for all L. Then,

$$\begin{aligned} \mathcal{L}^n (A_v) &= \int \chi_{A_v} \, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^{n-1} \times \mathbb{R}} \chi_{A_v} (x, y) \ d\mathcal{L}^n (x, y) \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{A_v} (x, y) \ d\mathcal{L}^1 (y) \right) \ d\mathcal{L}^{n-1} (x) \end{aligned}$$

; for inner integral x is fixed and hence $\chi_{A_v} (x, y) = 1$ if $(x, y) \in A_v$ and $\chi_{A_v} (x, y) = 0$ if $(x, y) \notin A_v$. Now let $(A_v)_x = \{y \in \mathbb{R}; (x, y) \in A_v\}$ then , $\chi_{(A_v)_x} = \begin{cases} 1 \ if \ y \in (A_v)_x \\ 0 \ if \ y \notin (A_v)_x \end{cases}$

, then $\chi_{A_v} (x, y) = \chi_{(A_v)_x} (y)$. This implies that

$$\begin{aligned} \mathcal{L}^n (A_v) &= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{(A_v)_x} (y) \ d\mathcal{L}^1 (y) \right) \ d\mathcal{L}^{n-1} (x) \\ &= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1 ((A_v)_x) \ d\mathcal{L}^{n-1} (x) . \end{aligned}$$

Notice that $\mathcal{L}^1$ is translation invariant; thus,

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} \mathcal{L}^1 ((A_v)_x) \ d\mathcal{L}^{n-1} (x) &= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1 (A_v \cap L_x) \ d\mathcal{L}^{n-1} (x) \\ &= \int_{\mathbb{R}^{n-1}} \mathcal{H}^1 (A_v \cap L_x) \ d\mathcal{H}^{n-1} (x) \\ &= 0. \end{aligned}$$

Finally we get,

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

This finishes the proof of **Claim 1** .

Claim # 2: gradient $f (x) = \left(\frac{\partial f}{\partial x_1} (x) , \cdots , \frac{\partial f}{\partial x_n} (x) \right)$ exists for $\mathcal{L}^n$ *a.e* x.

Proof of Claim # 2 : Applying **Claim 1** for $v = e_i = (0, \cdots , 1, \cdots)$ we get that $D_{e_i} f$ exists a.e; and hence for all $i = \{1, \cdots , n\}$ there exists E_i such that $\mu (E_i^c) = 0$, that is $D_{e_i} f$ exists on E_i. Let $E = \bigcup_{i=1}^{m} E_i^c$ then $\mu (E) = 0$. Moreover, for $(x_1, \cdots , x_n) \in E^c$, we have

$$\begin{aligned} D_{e_i} f (x_1, \cdots , x_n) &= \lim_{h \to 0} \frac{f ((x_1, \cdots , x_n) + h v_i) - f (x_1, \cdots , x_n)}{h} \\ &= \lim_{h \to 0} \frac{f ((x_1, \cdots , x_i + h, x_{i+1}, \cdots , x_n) + h v_i) - f (x_1, \cdots , x_n)}{h} \\ &= \frac{\partial f}{\partial x_i} (x_1, \cdots , x_n) . \end{aligned} \tag{5.0.4}$$

Claim #3: $D_v f(x) = v.grad\, f(x)$ for $\mathcal{L}^n$ $a.e\, x$.

Proof of Claim #3 : Let $\zeta \in C_c^\infty(\mathbb{R}^n)$ and let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$

$x \longrightarrow x + tv$.

Notice that T is one-to-one , then $|J_T| = \left| \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \right| = 1.$

Define $g(x) = f(x)\zeta(x - tv)$. Then by theorem 2.0.2 we get,

$$\begin{aligned} \int_{\mathbb{R}^n} g(x)\, d\mathcal{L}^n &= \int_{\mathbb{R}^n} f(x)\, \zeta(x - tv)\, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} g \circ T(x) \mid J_T \mid d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} g(x + tv)\, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} f(x + tv)\, \zeta(x)\, d\mathcal{L}^n. \end{aligned}$$

And so, $\int_{\mathbb{R}^n} \frac{f(x)\zeta(x - tv)}{t} d\mathcal{L}^n = \int_{\mathbb{R}^n} \frac{f(x + tv)\, \zeta(x)}{t} d\mathcal{L}^n$. This implies that,

$\int_{\mathbb{R}^n} \frac{f(x)\zeta(x - tv)}{t} d\mathcal{L}^n - \int_{\mathbb{R}^n} \frac{f(x)\, \zeta(x)}{t} d\mathcal{L}^n = \int_{\mathbb{R}^n} \frac{f(x + tv)\, \zeta(x)}{t} d\mathcal{L}^n - \int_{\mathbb{R}^n} \frac{f(x)\, \zeta(x)}{t} d\mathcal{L}^n.$

Hence, $\int_{\mathbb{R}^n} \frac{\zeta(x - tv) - \zeta(x)}{t} f(x)\, d\mathcal{L}^n = \int_{\mathbb{R}^n} \frac{f(x + tv) - f(x)}{t} \zeta(x)\, d\mathcal{L}^n$, which gives

$$-\int_{\mathbb{R}^n} \frac{\zeta(x) - \zeta(x - tv)}{t} f(x)\, d\mathcal{L}^n = \int_{\mathbb{R}^n} \frac{f(x + tv) - f(x)}{t} \zeta(x)\, d\mathcal{L}^n. \tag{5.0.5}$$

Now, applying theorem 3.0.10 on 5.0.5 and using the fact that $D_{e_i}\zeta(x_1, \cdots, x_n) = \frac{\partial \zeta}{\partial x_i}(x_1, \cdots, x_n)$; and by 5.0.4 that $D_{e_i} f(x_1, \cdots, x_n) = \frac{\partial f}{\partial x_i}(x_1, \cdots, x_n)$, we get

$$\int_{\mathbb{R}^n} D_v\zeta(x) f(x)\, d\mathcal{L}^n = \int_{\mathbb{R}^n} D_v f(x)\zeta(x)\, d\mathcal{L}^n \tag{5.0.6}$$

for all v. In fact for $v = e_i$ we get

$$\int_{\mathbb{R}^n} \frac{\partial \zeta}{\partial x_i} f\, d\mathcal{L}^n = \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i} \zeta\, d\mathcal{L}^n. \tag{5.0.7}$$

But since $\zeta(x)$ is $C_c^\infty(\mathbb{R}^n)$, we know that

$$D_v\zeta = \sum_{i=1}^n v_i \frac{\partial \zeta}{\partial x_i}, \tag{5.0.8}$$

where $v=\sum_{i=1}^{n} v_i e_i$. Using 5.0.6, 5.0.7 and 5.0.8 we get

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\,\zeta(x)\, d\mathcal{L}^n &= \int_{\mathbb{R}^n} D_v \zeta(x)\, f(x)\, d\mathcal{L}^n \\
&= \int_{\mathbb{R}^n} \sum_{i=1}^{n} v_i \frac{\partial \zeta}{\partial x_i} f\, d\mathcal{L}^n \\
&= \sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial \zeta}{\partial x_i} f\, d\mathcal{L}^n \\
&= \sum_{i=1}^{n} v_i \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i} \zeta\, d\mathcal{L}^n \\
&= \int_{\mathbb{R}^n} \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i} \zeta\, d\mathcal{L}^n \\
&= \int_{\mathbb{R}^n} v.grad\, f(x)\,\zeta(x)\, d\mathcal{L}^n.
\end{aligned}$$

hence $D_v f = v.grad\, f\ \mathcal{L}^n\ a.e.$

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

$$A_k = \left\{ x \in \mathbb{R}^n \ ;\ D_{v_k} f \; and \; grad\, f(x) \; exist \; and \; D_{v_k} f(x) = v_k.grad\, f(x) \right\} \; for \; k \in \mathbb{N}.$$

Define $A=\bigcap_{k=1}^{\infty} A_k$. Notice by **Claim # 2** that $\mathcal{L}^n(A_k^c)=0$, hence $\mathcal{L}^n(\mathbb{R}^n \setminus A_k)=0$ for all $k \in \mathbb{N}$.

$$\begin{aligned}
\mathcal{L}^n(\mathbb{R}^n \setminus A) &= \mathcal{L}^n\left(\mathbb{R}^n \setminus \left(\bigcap_{k=1}^{\infty} A_k\right)\right) \\
&= \mathcal{L}^n\left(\mathbb{R}^n \cap \left(\bigcap_{k=1}^{\infty} A_k\right)^c\right) \\
&= \mathcal{L}^n\left(\mathbb{R}^n \cap \left(\bigcup_{k=1}^{\infty} A_k^c\right)\right) \\
&= \mathcal{L}^n\left(\bigcup_{k=1}^{\infty} (\mathbb{R}^n \setminus A_k^c)\right) \\
&= \mathcal{L}^n\left(\bigcup_{k=1}^{\infty} (\mathbb{R}^n \setminus A_k)\right).
\end{aligned}$$

By countable subaddivity we get,

$$\mathcal{L}^n\left(\bigcup_{k=1}^{\infty}(\mathbb{R}^n\setminus A_k)\right) \leq \sum_{k=1}^{\infty}\mathcal{L}^n(\mathbb{R}^n\setminus A_k) = 0.$$

Thus, $\mathcal{L}^n(\mathbb{R}^n\setminus A)=0$.

Claim #4: f is differentiable at each point $x\in A$.

Proof of Claim #4: Fix an $x\in A$. Choose $v\in\partial B(0,1)\,,t\in\mathbb{R},t\neq 0$. Write

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

Then if $v'\in\partial B(0,1)\,,$ we get

$$\begin{aligned}|Q(x,v,t)-Q(x,v',t)| &= \left|\frac{f(x+tv)-f(x)}{t}-v.grad\,f(x)-\frac{f(x+tv')-f(x)}{t}+v'.grad\,f(x)\right| \\ &= \left|\frac{f(x+tv)-f(x+tv')}{t}+grad\,f(x)(v'-v)\right| \\ &\leq \left|\frac{f(x+tv)-f(x+tv')}{t}\right|+|grad\,f(x)(v'-v)| \end{aligned} \tag{5.0.9}$$

But

$$Lip(f)\geq\left|\frac{f(x+tv)-f(x+tv')}{(x+tv)-(x+tv)}\right|=\frac{|f(x+tv)-f(x+tv')|}{|t(v-v')|}. \tag{5.0.10}$$

Hence, replacing 5.0.10 in 5.0.9 we get

$$|Q(x,v,t)-Q(x,v',t)|\leq Lip(f)\,|v-v'|+|grad\,f(x)\,||v-v'|. \tag{5.0.11}$$

Let us show that

$$|grad\,f(x)\,|\leq\sqrt{n}\,Lip\,(f)\,. \tag{5.0.12}$$

Let $grad\,f(x)=\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_n}\right)$. Notice that

$$\left|\frac{\partial f}{\partial x_i}\right|=\left|\frac{f(x_1,\cdots,x_i+t,x_{i+1},\cdots,x_n)-f(x_1,\cdots,x_i,\cdots,x_n)}{t}\right|\leq Lip\,f(x)\,.$$

Hence,

$$\lim_{t\to 0}|\frac{f(x_1,\cdots,x_i+t,x_{i+1},\cdots,x_n)-f(x_1,\cdots,x_i,\cdots,x_n)}{t}|\leq Lip\,f(x)\,.$$

Thus,

$$\left|\frac{\partial f}{\partial x_i}\right|\leq Lip\,f.$$

Then,

$$\begin{aligned} |grad\, f|^2 &= |\frac{\partial f}{\partial x_1}|^2 + \cdots + |\frac{\partial f}{\partial x_n}|^2 \\ &\leq n\,(Lip\, f)^2 . \end{aligned}$$

Hence, $|grad\, f| \leq \sqrt{n}\, Lip\, f$, this proves 5.0.12. Replacing 5.0.12 in 5.0.11 we get

$$|Q\,(x, v, t) - Q\,(x, v', t)| \leq \left(\sqrt{n} + 1\right) Lip\,(f)\, |v - v'|. \tag{5.0.13}$$

Now fix $\epsilon > 0$ and choose N so large so that if $v \in \partial B\,(0, 1)$ then there exists $k \in \{1, \cdots, N\}$ and v_k such that

$$|v - v_k| \leq \frac{\epsilon}{2\,(\sqrt{n} + 1)\, Lip\,(f)}. \tag{5.0.14}$$

We want to show that $\lim\limits_{y \to x} \frac{|f\,(y) - f\,(x) - grad\, f\,(x) \,.\, (x - y)\,|}{|x - y|} = 0$. Replacing 5.0.14 in 5.0.13 for $v' = v_k$, we get

$$|Q\,(x, v, t) - Q\,(x, v_k, t)| < \frac{\epsilon}{2}. \tag{5.0.15}$$

Now by definition of v_k we have $\lim\limits_{t \to 0} Q\,(x, v_k, t) = 0$. This implies that for the chosen ϵ there exists δ such that if $|t| < \delta$ then $|Q\,(x, v_k, t)| < \frac{\epsilon}{2}$. Hence

$$|Q\,(x, v, t)| \leq |Q\,(x, v, t) - Q\,(x, v_k, t)| + |Q\,(x, v_k, t)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \tag{5.0.16}$$

Fix $\delta > 0$ and choose $y \in \mathbb{R}^n ; y \neq x$ and $|y - x| < \delta$. Write $v = \frac{y - x}{|y - x|}$ and hence $t = |x - y| < \delta$. Thus, using 5.0.16 we get

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

Hence, f is differentiable at x with $Df\,(x) = grad\, f\,(x)$.
We need to prove the theorem for the general case.
Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$
$\quad x \longrightarrow (f_1\,(x)\,, \cdots, f_m\,(x))$ be a Lipschitz function. Then, each
$f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$
$\quad x \longrightarrow f_i\,(x)$; is also Lipschitz since

$$\begin{aligned} |f_i\,(x) - f_i\,(y)\,| &\leq |f\,(x) - f\,(y)\,| \\ &\leq Lip\, f|x - y|. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{|f_i(x) - f_i(x)|}{|x-y|} &\leq Lip\, f \\ &= C. \end{aligned}$$

Thus, f_i is Lipschitz and by **Case 1** we get that f_i is differentiable a.e , which implies that f is differentiable a.e. □

Chapter 6

Linear maps and Jacobians

Definition 6.0.1.

1. A linear map $O : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is orthogonal if $(Ox) . (Oy) = x.y$ for all $x, y \in \mathbb{R}^n$.

2. A linear map $S : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is symmetric if $x. (Sy) = (Sx) .y$ for all $x, y \in \mathbb{R}^n$.

3. A linear map $D : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is diagonal if there exists $d_1, \cdots, d_n \in \mathbb{R}$ such that $D_x = (d_1 x_1, \cdots, d_n x_n)$ for all $x \in \mathbb{R}^n$.

4. Let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be linear. The adjoint of A is the linear map $A^* : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ defined by $x. (A^* y) = (Ax) .y$ for all $x \in \mathbb{R}^n, y \in \mathbb{R}^m$.

Theorem 6.0.2. *Properties of Linear maps*

1. $A^{**} = A$.

2. $(A \circ B)^* = B^* \circ A^*$.

3. $O^* = O^{-1}$, *if* $O : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *is orthogonal.*

4. $S^* = S$ *if* $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *is symmetric.*

5. *If* $O : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *is orthogonal then* $n \leq m$ *and* $O^* \circ O = I$ *on* $\mathbb{R}^n$ *and* $O \circ O^* = I$ *on* $O(\mathbb{R}^n)$.

Theorem 6.0.3. *Polar decomposition.*
Let $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a linear map.*

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

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

Proof. 1. Consider $C = L^* \circ L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Then, by the definition of the adjoint

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

Also $(Cx).x = Lx.Lx \geq 0$, hence C is symmetric, non negative definite. Then there exist $\mu_1, \cdots, \mu_n \geq 0$ and an orthogonal basis $\{x_k\}_{k=1}^n$ of $\mathbb{R}^n$ such that

$$Cx_k = \mu_k x_k \ \ (k = 1, \cdots, n).$$

Write $\mu_k = {\lambda_k}^2$; $\lambda_k \geq 0 \ \ (k = 1, \cdots, n)$.
Claim: There exists an orthonormal set $\{z_k\}_{k=1}^n$ in $\mathbb{R}^m$ such that $Lx_k = \lambda_k z_k \, for \, k = \{1, \cdots, n\}$.
Proof of Claim: **Case 1:**If $\lambda_k \neq 0$, define $z_k = \dfrac{1}{\lambda_k} Lx_k$. Then, if $\lambda_k, \lambda_l \neq 0$ we get

$$\begin{aligned}z_k.z_l &= \frac{1}{\lambda_k \lambda_l} Lx_k Lx_l \\ &= \frac{1}{\lambda_k \lambda_l}(Cx_k).x_l \\ &= \frac{1}{\lambda_k \lambda_l}{\lambda_k}^2 x_k.x_l \\ &= \frac{\lambda_k}{\lambda_l}\delta_{kl}\end{aligned}$$

where $\delta_{kl} = \begin{cases} 1 \, if \ k = l \\ 0 \, if \ k \neq l \end{cases}$. Thus the set $\{z_k \, ; \lambda_k \neq 0\}$ is orthonormal.

Case 2 : If $\lambda_k = 0$ then ${\lambda_k}^2 = 0$; this implies that $\mu_k = 0$. but, $Cx_k = \mu_k x_k$ then, $Cx_k = 0$. So $L^* \circ L(x_k) = 0$, hence $(L^* \circ L(x_k)).x_k = 0$. And by the definition of the adjoint we get $L(x_k).L(x_k) = 0$ which implies $|L(x_k)|^2 = 0$ thus, $L(x_k) = 0$. Define $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by $Sx_k = \lambda_k x_k \ \ (k = 1, \cdots, n)$
and
$O : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $Ox_k = z_k$. Then,

$$\begin{aligned}O \circ Sx_k &= O(\lambda_k x_k) \\ &= \lambda_k O x_k \\ &= \lambda_k z_k \\ &= Lx_k.\end{aligned}$$

Hence, $L = O \circ S$. Rest to show that S is symmetric and O is orthogonal.

- S **is symmetric.** To see this , let $x, y \in \mathbb{R}^n$ where $x = \sum_{k=1}^n \alpha_k x_k$ and $y = \sum_{l=1}^n \beta_l x_l$.

Then,

$$\begin{aligned}
x.S(y) &= \left(\sum_{k=1}^{n} \alpha_k x_k\right).S\left(\sum_{l=1}^{n} \beta_l x_l\right) \\
&= \sum_{k=1}^{n} \alpha_k x_k . \left(\sum_{l=1}^{n} \beta_l S(x_l)\right) \\
&= \sum_{k,l=1}^{n} \alpha_k \beta_l x_k . S(x_l) \\
&= \sum_{k,l=1}^{n} \alpha_k S(x_k) . x_l \\
&= \sum_{k=1}^{n} \alpha_k S(x_k) . \sum_{l=1}^{n} \beta_l S(x_l) \\
&= S\left(\sum_{k=1}^{n} \alpha_k x_k\right) . \sum_{l=1}^{n} \beta_l x_l \\
&= S(x).y.
\end{aligned}$$

- O **is orthogonal.** To see this , let

$$\begin{aligned}
O_{x_k}.O_{x_l} &= z_k.z_l \\
&= \delta_{kl} \\
&= x_k.x_l.
\end{aligned}$$

For any $x, y \in \mathbb{R}^n$, let $x = \sum_{k=1}^{n} \alpha_k x_k$ and $y = \sum_{l=1}^{n} \beta_l x_l$. Then,

$$\begin{aligned}
Ox.Oy &= O\left(\sum_{k=1}^{n} \alpha_k x_k\right).O\left(\sum_{l=1}^{n} \beta_l x_l\right) \\
&= \sum_{k=1}^{n} \alpha_k (x_k) . \sum_{l=1}^{n} \beta_l O(x_l) \\
&= \sum_{k=1}^{n} \alpha_k \sum_{l=1}^{n} \beta_l O(x_k) O(x_l) \\
&= \sum_{k=1}^{n} \alpha_k \sum_{l=1}^{n} \beta_l (x_k)(x_l) \\
&= \left(\sum_{k=1}^{n} \alpha_k x_k\right) . \left(\sum_{l=1}^{n} \beta_l x_l\right) \\
&= x.y.
\end{aligned}$$

2. For the case where $n \geq m$, let $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, then $L^* : \mathbb{R}^m \longrightarrow \mathbb{R}^n$. But $L^* = O \circ S$ such that O is orthogonal and S is symmetric, then

$$\begin{aligned} L &= (L^*)^* \\ &= (O \circ S)^* \\ &= S^* \circ O^* \\ &= S \circ O^*. \end{aligned}$$

□

Definition 6.0.4. Let $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map. If $n \leq m$, write $L = O \circ S$ and define the Jacobian of L to be $[[L]] = |det\, S|$.
Note that $[[L]] = [[L^*]]$.

Theorem 6.0.5. *For* $n \leq m$; $[[L]]^2 = det\,(L^* \circ L)$.

Proof. Write $L = O \circ S$ and $L^* = S^* \circ O^* = S \circ O^*$. Then,

$$\begin{aligned} L^* \circ L &= S \circ O^* \circ O \circ S \\ &= S^2\,(O^* \circ O). \end{aligned}$$

Hence,

$$\begin{aligned} det\,(L^* \circ L) &= det\,(S^2) \\ &= det\,(S.S) \\ &= (det\,S)^2 \\ &= [[L]]^2. \end{aligned}$$

□

Definition 6.0.6. 1. For $n \leq m$; define $\Lambda\,(m,n) = \left\{\lambda : \{1, \cdots, n\} \rightarrow \{1, \cdots, m\};\, \lambda\, is\, increasing\right\}$.

2. For each $\lambda \in \Lambda\,(m,n)$; define $P_\lambda : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ by $P_\lambda\,(x_1, \cdots, x_m) = \left(x_{\lambda(1)}, \cdots, x_{\lambda(n)}\right)$.

Remark 6.0.7. For each $\lambda \in \Lambda\,(m,n)$, there exists an n-dimensional subspace $S_\lambda = span\left\{e_{\lambda(1)}, \cdots, e_{\lambda(n)}\right\} \in \mathbb{R}^m$ such that P_λ is the projection of $\mathbb{R}^m$ onto S_λ.

Theorem 6.0.8. *Binet-Cauchy Formula*
Let $n \leq m$, $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a linear map. Then,*

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

Notice that $P_\lambda \circ L : \mathbb{R}^n \longrightarrow \mathbb{R}^m \longrightarrow \mathbb{R}^n$ *and*

$$P_\lambda \circ L = \left(det\,(P_{\lambda_1} \circ L)\right)^2 + \left(det\,(P_{\lambda_2} \circ L)\right)^2 + \cdots.$$

Remark 6.0.9. : In order to calculate $[[L]]^2$, we compute the sums of the squares of the determinants of each $(n \times n)$ submatrix of the $(m \times n)$ matrix representing L.

Chapter 7

The Area Formula

In this section, we will show that if $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a Lipschitz function such that $n \leq m$. Then for each $\mathcal{L}^n$ measurable set $A \subset \mathbb{R}^n$

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

Notice that the left hand side of this equation gives the area of $A \subset \mathbb{R}^n$.

Lemma 7.0.1. *If $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear map such that $n \leq m$, then*

$$\mathcal{H}^n (L (A)) = [[L]] \mathcal{L}^n (A) ; \forall A \subset \mathbb{R}^n.$$

Proof. 1. Let $L = O \circ S$, then $[[L]] = |det \, S|$.

Case 1: If $[[L]] = 0$. Since $L = O \circ S$ and $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear symmetry, then $n = dim \, Ker \, S + dim \, Im \, S$. But $[[L]] = 0$, thus $|det \, S| = 0$ and hence S is not invertible. It follows that S is not one-to-one , thus $Ker \, S \neq \{0\}$,implying that $dim \, Ker \, S \geq 1$. Finally we get $dim \, Im \, S \leq n-1$. Hence $dim \, S (\mathbb{R}^n) \leq n-1$ and hence $dim \, L (\mathbb{R}^n) \leq n-1$. Using the fact that $dim \, L (\mathbb{R}^n) \leq n-1$ we get $\mathcal{H}^n (L (\mathbb{R}^n)) = 0$.

Case 2: If $[[L]] > 0$. Notice that

$$\begin{aligned} \frac{\mathcal{H}^n (L (B (x, r)))}{\mathcal{L}^n (B (x, r))} &= \frac{\mathcal{L}^n (O^* \circ L (B (x, r)))}{\mathcal{L}^n (B (x, r))} \\ &= \frac{\mathcal{L}^n (O^* \circ O \circ S (B (x, r)))}{\mathcal{L}^n (B (x, r))} \\ &= \frac{\mathcal{L}^n (S (B (x, r)))}{\mathcal{L}^n (B (x, r))} \\ &= \frac{\mathcal{L}^n (S (B (0, 1)))}{\alpha (n)}. \end{aligned} \tag{7.0.1}$$

But using the change of variables formula for $\mathcal{L}^n$ (see theorem 2.0.2) we get

$$\frac{\mathcal{L}^n (S (B (0, 1)))}{\alpha (n)} = |det \, S| = [[L]]. \tag{7.0.2}$$

Plugging 7.0.2 in 7.0.1 we get

$$\frac{\mathcal{H}^n\left(L\left(B\left(x,r\right)\right)\right)}{\mathcal{L}^n\left(B\left(x,r\right)\right)} = [[L]]. \tag{7.0.3}$$

Notice that the Jacobian of S is equal to the determinant of S which is a number.

2. Define $v\left(A\right) = \mathcal{H}^n\left(L\left(A\right)\right)$ for all $A \subset \mathbb{R}^n$. We will prove that v is a radon measure and is absolutely continuous with respect to $\mathcal{L}^n$. First let us prove that v is a measure.
a) $v\left(\phi\right) = \mathcal{H}^n\left(L\left(\phi\right)\right) = 0$.
b)

$$\begin{aligned} v\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathcal{H}^n\left(L\left(\bigcup_{n=1}^{\infty} A_n\right)\right) \\ &= \mathcal{H}^n\left(\bigcup_{n=1}^{\infty}\left(L\left(A_n\right)\right)\right) \\ &\leq \sum_{n=1}^{\infty} \mathcal{H}^n\left(L\left(A_n\right)\right) \\ &= \sum_{n=1}^{\infty} v\left(A_n\right). \end{aligned}$$

Hence, v is a measure.
Next we will prove that v is borel regular.
Let $A \subseteq \mathbb{R}^n$ then $L\left(A\right) \subseteq L\left(\mathbb{R}^n\right)$. $\mathcal{H}^n$ is borel regular then there exists a borel set C such that $L\left(A\right) \subset C$ and such that

$$\mathcal{H}^n\left(L\left(A\right)\right) = \mathcal{H}^n\left(C\right). \tag{7.0.4}$$

Also since $\mathcal{H}^n$ is borel regular then there exists a borel set B such that $\mathcal{H}^n\left(A\right) = \mathcal{H}^n\left(B\right)$. Notice that since $A \subseteq B$ then $L\left(A\right) \subseteq L\left(B\right) \subseteq L\left(\mathbb{R}^n\right)$. Take $C \cap L\left(B\right)$ and let $D := C \cap L\left(B\right) \subseteq L\left(\mathbb{R}^n\right)$. Since L is bijective then there exists a set E such that $L\left(E\right) = D$ that is $E = L^{-1}\left(D\right)$. $C \cap L\left(B\right) \subseteq C$ and $L\left(A\right) \subseteq C \cap L\left(B\right)$ then

$$\mathcal{H}^n\left(C \cap L\left(B\right)\right) \leq \mathcal{H}^n\left(C\right) = \mathcal{H}^n\left(L\left(A\right)\right) \leq \mathcal{H}^n\left(C \cap L\left(B\right)\right).$$

Hence,

$$\mathcal{H}^n\left(L\left(A\right)\right) = \mathcal{H}^n\left(C \cap L\left(B\right)\right) = \mathcal{H}^n\left(L\left(E\right)\right).$$

Thus,

$$v\left(A\right) = v\left(E\right),$$

and v is a radon measure.

Next we will prove that $v \ll \mathcal{L}^n$.
Let $A \subset \mathbb{R}^n$ such that $\mathcal{L}^n(A) = 0$. We want to prove that $v(A) = 0$. But, $\mathcal{L}^n(A) = \mathcal{H}^n(A) = 0$, hence $\mathcal{H}^n(L(A)) = 0$. This implies that $v(A) = 0$.
Now recalling the definition of $\mathcal{D}_{\mathcal{L}^n} v$ (see theorem 3.0.1) , we have

$$\begin{aligned} \mathcal{D}_{\mathcal{L}^n} v(x) &= \lim_{t\to 0} \frac{v(B(x,r))}{\mathcal{L}^n(B(x,r))} \\ &= \lim_{t\to 0} \frac{\mathcal{H}^n(L(B(x,r)))}{\mathcal{L}^n(B(x,r))} \\ &= [[L]]. \end{aligned}$$

Where last step comes from 7.0.3. Hence for all borel sets $B \subset \mathbb{R}^n$ we have

$$\begin{aligned} v(B) &= \mathcal{H}^n(L(B)) \\ &= \int_B \mathcal{D}_{\mathcal{L}^n} v(B)\, d\mathcal{L}^n \\ &= \int_B [[L]]\, d\mathcal{L}^n \\ &= [[L]]\mathcal{L}^n(B). \end{aligned}$$

Thus,

$$\mathcal{H}^n(L(B)) = [[L]]\mathcal{L}^n(B). \tag{7.0.5}$$

We still need to show that

$$\mathcal{H}^n(L(A)) = [[L]]\mathcal{L}^n(A) \text{ ; for any set } A \subset \mathbb{R}^n.$$

To see that, let $A \subset \mathbb{R}^n$. Since v is borel then there exists a set B_1 such that $A \subseteq B_1$ and $v(A) = v(B_1) = v(B)$.
Also since $\mathcal{L}^n$ is borel then there exists a set B_2 such that $A \subseteq B_2$ and $\mathcal{L}^n(A) = \mathcal{L}^n(B_2) = \mathcal{L}^n(B)$. Notice that $B = B_1 \cap B_2$ then $A \subseteq B \subseteq B_1$ and hence we get

$$v(A) \le v(B) \le v(B_1) = v(A).$$

On the other hand $A \subseteq B \subseteq B_2$ then

$$\mathcal{L}^n(A) \le \mathcal{L}^n(B) \le \mathcal{L}^n(B_2) = \mathcal{L}^n(A).$$

Thus,

$$v(A) = v(B) = [[L]]\mathcal{L}^n(B) = [[L]]\mathcal{L}^n(A).$$

□

Lemma 7.0.2. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a Lipschitz function. If $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$ measurable then :

1. *$f(A)$ is $\mathcal{H}^n$ measurable.*

2. The multiplicity function from y to $\mathcal{H}^0\left(A \cap f^{-1}\{y\}\right)$ is $\mathcal{H}^n$ measurable on $\mathbb{R}^m$.

3. $\displaystyle\int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n \leq (Lip\, f)^n\, \mathcal{L}^n(A)$.

Proof. 1. Assume A is bounded , then for all $i \in \mathbb{N}$ there exists compact sets $K_i \subset A$ such that $\mathcal{L}^n(K_i) \geq \mathcal{L}^n(A) - \frac{1}{i}$. And hence, $\mathcal{L}^n(A) - \mathcal{L}^n(K_i) \leq \frac{1}{i}$. But $\mathcal{L}^n(A \setminus K_i) = \mathcal{L}^n(A) - \mathcal{L}^n(K_i)$ and thus, $\mathcal{L}^n(A \setminus K_i) \leq \frac{1}{i}$. Notice

$$\mathcal{L}^n\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right) = 0. \tag{7.0.6}$$

Moreover, since f is a continuous function then $f(K_i)$ is compact and thus $\mathcal{H}^n$ measurable. So, $f\left(\bigcup_{i=1}^{\infty} K_i\right) = \bigcup_{i=1}^{\infty} f(K_i)$ is $\mathcal{H}^n$ measurable. Let us show that

$$\mathcal{H}^n\left(f(A) - f\left(\bigcup_{i=1}^{\infty} K_i\right)\right) \leq \mathcal{H}^n\left(f\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right)\right).$$

To see this, we have that $f(A) - f\left(\bigcup_{i=1}^{\infty} K_i\right) = f(A) \cap f\left(\bigcup_{i=1}^{\infty} K_i\right)^c$; but

$$f(A) \cap f\left(\bigcup_{i=1}^{\infty} K_i\right)^c \subset f\left(A \cap \left(\bigcup_{i=1}^{\infty} K_i\right)^c\right) = f\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right).$$

This implies that

$$\mathcal{H}^n\left(f(A) - f\left(\bigcup_{i=1}^{\infty} K_i\right)\right) \leq \mathcal{H}^n\left(f\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right)\right). \tag{7.0.7}$$

Notice that by Theorem 4.2.3, $\mathcal{H}^n\left(f\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right)\right) \leq (Lip\, f)^n\, \mathcal{L}^n\left(A \setminus \bigcup_{i=1}^{\infty} K_i\right)$ and hence replacing in 7.0.7 and using 7.0.6, we get

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

Which implies that $f(A)$ is $\mathcal{H}^n$ measurable.

2. Fix $k \in \mathbb{N}$. Let $B_k = \left\{Q; Q = (a_1, b_1) \times \cdots \times (a_n, b_n)\,;\ a_i = \frac{c_i}{k}, b_i = \frac{c_i + 1}{k}, c_i \text{ are integers}, i = 1, 2, \cdots, n\right\}$. Notice that $\mathbb{R}^n = \bigcup_{Q \in B_k} Q$. Let $g_k = \sum_{Q \in B_k} \chi_{f(A \cap Q)}$, then g_k is $\mathcal{H}^n$ measurable, since $A \cap Q$ is measurable. Notice that $g_k(y)$ is equal to the number of

cubes $Q \in B_k$ such that $f^{-1}\{y\} \cap (A \cap Q) \neq \phi$; let us show that $g_k(y)$ converges to $\mathcal{H}^0(A \cap f^{-1}\{y\})$ as $k \to \infty$ for each $y \in \mathbb{R}^m$. Since, let $g_k = \sum_{Q \in B_k} \chi_{f(A\cap Q)}$, then

$$\begin{aligned}
\lim_{k\to\infty} g_k &= \lim_{k\to\infty} \sum_{Q\in B_k} \chi_{f(A\cap Q)} \\
&= \sum_{Q\in\bigcup_{k=1}^{\infty} B_k} \chi_{f(A\cap Q)}.
\end{aligned}$$

Then,

$$\begin{aligned}
\lim_{k\to\infty} g_k(y) &= \sum_{Q\in\mathbb{R}^n} \chi_{f(A\cap Q)}(y) \\
&= \sum_{x\in f^{-1}\{y\}} \chi_{f(A\cap Q)} \\
&= \mathcal{H}^0\left(A\cap f^{-1}\{y\}\right). \qquad (7.0.8)
\end{aligned}$$

So $g : y \longrightarrow \mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$ measurable.

3. Using the Monotone convergence theorem (see Theorem 3.0.8) and 7.0.8 we get,

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

□

Lemma 7.0.3. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Let $t > 1$ and $B = \left\{x\,; Df(x) \text{ exists}, Jf(x) > 0\right\}$. Then there exists a countable collection $\{E_k\}_{k=1}^{\infty}$ of borel subsets of $\mathbb{R}^n$ such that:

1. $B = \bigcup_{k=1}^{\infty} E_k$

2. $f \restriction_{E_k}$ *is one-to-one for* $k \in \mathbb{N}$

3. For each $k \in \mathbb{N}$, there exists a symmetric automorphism $T_k : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that :

- *$Lip\left(\left(f \restriction_{E_k}\right) \circ T_k^{-1}\right) \leq t$*

- *$Lip\left(T_k \circ \left(f \restriction_{E_k}\right)^{-1}\right) \leq t$*

- *$t^{-n}|det\, T_k| \leq Jf \restriction_{E_k} \leq t^n |det\, T_k|$.*

Proof. 1. Fix $\epsilon > 0$ so that $\frac{1}{t} + \epsilon < 1 < t - \epsilon$. Let $B \subset \mathbb{R}^n$. Since $\mathbb{R}^n$ is separable, let C be a countable dense subset of B. Since any set of symmetric automorphism on $\mathbb{R}^n$ is isomorphic to $\mathbb{R}^{\frac{n(n+1)}{2}}$, we have a countable dense subset $\mathcal{S}$ of symmetric automorphism T on $\mathbb{R}^n$, with operator norm $\|T\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|T(x)|}{|x|}$. Note that for all $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ a symmetric automorphism T is Lipschitz.To see that notice that T is a linear function , hence continous. Thus since all norms are equivalent in a finite dimensional space $\mathbb{R}^{n\times n}$, continuity is equivalent to boundedness. And we have

$$|\frac{T(x) - T(y)}{x - y}| = |\frac{T(x - y)}{x - y}| \leq \|T\|$$

so, $|T(x) - T(y)| \leq \|T\| |x - y|$. Thus,

$$Lip\, T \leq \|T\|. \tag{7.0.9}$$

Define $E(c, T, i)$, where $c \in C$, $T \in \mathcal{S}$ and $i \in \mathbb{N}$, to be the set of all $b \in B \cap B\left(c, \frac{1}{i}\right)$ that satisfies

$$\left(\frac{1}{t} + \epsilon\right)|Tv| \leq |Df(b)\, v| \leq (t - \epsilon)\,|Tv| \text{ for all } v \in \mathbb{R}^n \tag{7.0.10}$$

and

$$|f(a) - f(b) - Df(b)(a - b)| \leq \epsilon |T(a - b)| \text{ for all } a \in B\left(b, \frac{2}{i}\right). \tag{7.0.11}$$

Notice that $E(c, T, i)$ is a borel set since Df is borel measurable. Letting $v = a - b$, we get

$$\frac{1}{t}|T(a - b)| \leq |f(a) - f(b)| \leq t|T(a - b)| \; for\, b \in E(c, T, i) \;\; , a \in B\left(b, \frac{2}{i}\right) \tag{7.0.12}$$

Claim: If $b \in E(c, T, i)$ then

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

Proof of Claim: Write $Df(b) = L = O \circ S$. Then,

$$Jf(b) = [[Df(b)]] = |det\, S|. \tag{7.0.13}$$

Moreover for all $v' \in \mathbb{R}^n$ we have

$$|Df(b)\, v'| = |O \circ S(v')| = |S(v')| \tag{7.0.14}$$

Replacing 7.0.14 in 7.0.10 for $v' = T^{-1}(v)$, we get,

$$\left(\frac{1}{t} + \epsilon\right) |\,(v)\,| \leq |S \circ T^{-1}(v)\,| \leq (t - \epsilon)\,|v| \text{ for all } v \in \mathbb{R}^n. \tag{7.0.15}$$

Thus, $\left(S \circ T^{-1}\right)(B(0,1)) \subset B(0, t - \epsilon)$. This gives $\mathcal{L}^n\Big(\left(S \circ T^{-1}\right)(B(0,1))\Big) \leq \mathcal{L}^n(B(0, t - \epsilon))$.
But $\mathcal{L}^n\Big(\left(S \circ T^{-1}\right)(B(0,1))\Big) = det|S \circ T^{-1}|\alpha(n)$ (see Theorem 2.0.2).Thus,

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

That is,

$$|det\, S| \leq (t - \epsilon)^n\, |det\, T|. \tag{7.0.16}$$

Plugging 7.0.13 in 7.0.16, we get $Jf(b) \leq (t - \epsilon)^n\, |det\, T|$. This proves the right hand side of our **Claim**. Now in order to prove the other inequality we notice that by 7.0.15 we have that $B\left(0, \frac{1}{t} + \epsilon\right) \subset \left(S \circ T^{-1}\right)(B(0,1))$. Hence,

$$\begin{aligned} \mathcal{L}^n\left(B\left(0, \frac{1}{t} + \epsilon\right)\right) &\leq \mathcal{L}^n\left(\left(S \circ T^{-1}\right)(B(0,1))\right) \\ &= det|S \circ T^{-1}|\,(\alpha(n)). \end{aligned}$$

This implies that

$$\alpha(n)\left(\frac{1}{t} + \epsilon\right)^n \leq det|S \circ T^{-1}|\,(\alpha(n)) \tag{7.0.17}$$

but,

$$det|S \circ T^{-1}| = |det\, S|\,\left|detT^{-1}\right| = |det\, S|\,.\frac{1}{|det\, S|}. \tag{7.0.18}$$

Thus replacing 7.0.18 in 7.0.17 we get

$$|det\, S| \geq \left(\frac{1}{t} + \epsilon\right)^n |det\, T|\,. \tag{7.0.19}$$

Plugging 7.0.13 back in 7.0.19, we get the inequality we want, and the claim is proved. Let $\{E_k\}_{k=1}^{\infty} = \left\{E(c, T, i) \ ; \ c \in C \ , T \in S \ \ i \in \mathbb{N}\right\}$. Fix $b \in B$. First, we will show that there exists $T \in \mathcal{S}$ such that

$$Lip\left(T \circ S^{-1}\right) \leq \left(\frac{1}{t} + \epsilon\right)^{-1} \tag{7.0.20}$$

and

$$Lip\left(S \circ T^{-1}\right) \leq t - \epsilon. \tag{7.0.21}$$

Since S is symmetric automorphism , then for any $\epsilon' > 0$ there exists $T \in \mathcal{S}$ such that $\|T - S\| < \epsilon'$. This implies that

$$\|\left(T \circ S^{-1} - Id\right) \circ S\| < \epsilon'.$$

Thus, for all $x \in \mathbb{R}^n$ we have $\dfrac{|\left(T \circ S^{-1} - Id\right) \circ S(x)|}{|x|} < \epsilon'$. But since S is bijective, for all $y \in \mathbb{R}^n$ there exists an $x \in \mathbb{R}^n$ such that $x = S^{-1}(y)$. Hence,

$$\frac{|\left(T \circ S^{-1} - Id\right) \circ S\left(S^{-1}(y)\right)|}{|S^{-1}(y)|} < \epsilon',$$

that is,

$$|\left(T \circ S^{-1} - Id\right)(y)| < \epsilon' |S^{-1}(y)| < \epsilon' \|S^{-1}\| |y|.$$

If we divide both sides by $|y|$ we get,

$$\frac{|\left(T \circ S^{-1} - Id\right)(y)|}{|y|} < \epsilon' \|S^{-1}\| for\, all\, y \in \mathbb{R}^n.$$

So, $\|T \circ S^{-1} - Id\| < \epsilon' \|S^{-1}\|$, which implies $\|T \circ S^{-1}\| < 1 + \epsilon' \|S^{-1}\|$. Thus,

$$Lip\left(T \circ S^{-1}\right) \leq 1 + \epsilon' \|S^{-1}\|. \tag{7.0.22}$$

We want

$$1 + \epsilon' \|S^{-1}\| = \left(\frac{1}{t} + \epsilon\right)^{-1} \tag{7.0.23}$$

, that is we want , $1 + \epsilon' \|S^{-1}\| = \dfrac{1}{\frac{1}{t} + \epsilon}$. This means,

$$\left(\frac{1}{t} + \epsilon\right)\left(1 + \epsilon' \|S^{-1}\|\right) = 1.$$

Hence,

$$\frac{1}{t} + \epsilon' \frac{\|S^{-1}\|}{t} + \epsilon + \epsilon\epsilon' \|S^{-1}\| = 1;$$

which implies

$$\epsilon^{'}\left(\frac{\|S^{-1}\|}{t}+\epsilon\|S^{-1}\|t\right)=1-\frac{1}{t}-\epsilon$$

Thus, for

$$\epsilon^{'}=\frac{1-\frac{1}{t}-\epsilon}{\left(\frac{\|S^{-1}\|}{t}+\epsilon\|S^{-1}\|t\right)},$$

we have 7.0.23. Replacing 7.0.23 in 7.0.22, we get 7.0.20. similiar work gives us 7.0.21. Next, let us show $b\in E\left(c,T,i\right)$. First we show that b satisfies 7.0.11. Since

$$\lim_{a\to b}\frac{|f\left(a\right)-f\left(b\right)-Df\left(b\right)\left(a-b\right)|}{|a-b|}=0.$$

Then for $\frac{\epsilon}{Lip\left(T^{-1}\right)}$ there exits δ, such that if $|a-b|<\delta$ we have

$$|f\left(a\right)-f\left(b\right)|<\frac{\epsilon}{Lip\left(T^{-1}\right)}. \tag{7.0.24}$$

Choose i such that $\frac{2}{i}<\delta$, then for all $a\in B\left(b,\frac{2}{i}\right)$ we get

$$\begin{aligned}|f\left(a\right)-f\left(b\right)-Df\left(b\right)\left(a-b\right)| &\leq \frac{\epsilon}{Lip\left(T^{-1}\right)}|a-b|\\ &= \frac{\epsilon}{Lip\left(T^{-1}\right)}|T^{-1}\left(T\left(a\right)\right)-T^{-1}\left(T\left(b\right)\right)|\\ &\leq \frac{\epsilon}{Lip\left(T^{-1}\right)}Lip\left(T^{-1}\right)|T\left(a\right)-T\left(b\right)|\\ &= \epsilon|T\left(a-b\right)|.\end{aligned}$$

Choosing $c\in C$ such that $|b-c|<\frac{1}{i}$ (we can because C is dense in B). This shows that b satisfies 7.0.11. Rest to show that b satisfies 7.0.12. Since $Df\left(b\right)=L=O\circ S$ then, for all $v\in\mathbb{R}^{n}$

$$\begin{aligned}|Df\left(b\right)\left(v\right)| &= |O\circ S\left(v\right)|\\ &= |S\left(v\right)|\\ &= |S\circ T^{-1}\circ T\left(v\right)|\\ &= |S\circ T^{-1}\left(T\left(v\right)\right)|\\ &= |S\circ T^{-1}\left(T\left(v\right)\right)-S\circ T^{-1}\left(T\left(0\right)\right)|\\ &\leq Lip|S\circ T^{-1}||Tv|\\ &\leq t-\epsilon\,|T\left(v\right)|.\end{aligned}$$

where the last inequality comes from 7.0.21. Also,

$$\begin{aligned} |T(v)| &= |T \circ S^{-1} \circ S(v)| \\ &= |T \circ S^{-1}(S(v))| \\ &= |T \circ S^{-1}(S(v)) - T \circ S^{-1}(S(0)| \\ &\leq Lip\left(T \circ S^{-1}\right) |S(v)|. \end{aligned}$$

This implies that

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

Where the last inequality comes from 7.0.20. This shows that b satisfies 7.0.12. So,

$$\begin{aligned} |S(v)| &= |Df(b)(v)| \\ &\geq \left(\frac{1}{t} + \epsilon\right) |T(v)|. \end{aligned}$$

As this conclusion holds for all $b \in B$ then $B = \bigcup_{k=1}^{\infty} E_k$.

2. Choose any set E_k which is of the form $E(c, T, i)$ for some $c \in C$,$T \in S$,$i \in \mathbb{N}$. Let $T_k = T$. Using (7.0.12) we get

$$\frac{1}{t}|T_k(a-b)| \leq |f(a) - f(b)| \leq t|T_k(a-b)| \; for\, all\, a, b \in E_k. \qquad (7.0.25)$$

Let us show that $f \restriction_{E_k}$ is one-to-one.If $f(a) = f(b)$,then let $\frac{1}{t}|T_k(a-b)| \leq 0$, hence $\frac{1}{t}|T_k(a-b)| = 0$, which implies that $T_k(a-b) = 0$, and hence $a = b$ (because T is a symmetric automorphism).
Let $T_k^{-1}(x) = a$ and $T_k^{-1}(y) = b$, then using 7.0.25, we get that

$$|\frac{1}{t}T_k\left(T_k^{-1}(x) - T_k^{-1}(y)\right)| \leq |f \restriction_{E_k} \left(T_k^{-1}(x) - f \restriction_{E_k} T_k^{-1}(y)\right| \leq |t\, T_k T_k^{-1}(x) - T_k^{-1}(y))|$$

and hence,

$$\frac{1}{t}|x-y| \leq |f \circ T_k^{-1}(x) - f \circ T_k^{-1}(y)| \leq t|x-y|.$$

Thus,

$$\frac{1}{t} \leq \frac{|f \circ T_k^{-1}(x) - f \circ T_k^{-1}(y)|}{|x-y|} \leq t.$$

Taking the supremum on both sides we get

$$Lip\left((f \restriction_{E_k}) \circ T_k^{-1}\right) \leq t$$

and

$$Lip\left(T_k \circ (f\restriction_{E_k})^{-1}\right) \le t.$$

Finally, notice that the Claim gives us the estimate

$$t^{-n}|det\, T_k| \le Jf\restriction_{E_k} \le t^n|det\, T_k|.$$

□

Theorem 7.0.4. *The Area Formula*
Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a Lipschitz function such that* $n \le m$*. Then for each* $\mathcal{L}^n$ *measurable set* $A \subset \mathbb{R}^n$

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

Proof. Using Rademacher's theorem (see Theorem 5.0.5), we may assume that $Df(x)$ and $Jf(x)$ exist for all $A \subset \mathbb{R}^n$ and $\mathcal{L}^n(A) < \infty$. There are 2 cases to be considered :
Case 1: $A \subset \{Jf(x) > 0\}$.
Fix $k > 0$ and $t > 1$. Choose borel sets $\{E_j\}_{j=1}^{\infty}$ as in Lemma 7.0.12 assuming that they are disjoint. Define

$$B_k = \left\{Q\,; Q = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)\ \ , a_i = \frac{c_i}{k}, b_i = \frac{c_i + 1}{k}, c_i\, integers\ , i = 1, 2 \cdots, n\right\}.$$

Set $F_j^i = E_j \cup Q_i \cup A; (Q_i \in B_k\,, j \in \mathbb{N})$, then the sets F_j^i are disjoint and $A = \bigcup_{i,j=1}^{\infty} F_j^i$. To see this, let

$$\begin{aligned}
\bigcup_{i,j=1}^{\infty} F_j^i &= \bigcup_{i,j=1}^{\infty} (E_j \cup Q_i \cup A) \\
&= A \cap \left(\bigcup_{i,j=1}^{\infty} E_j \cup Q_i\right) \\
&= A \cap \left(\bigcup_{j=1}^{\infty} E_j \cup \bigcup_{i=1}^{\infty} Q_i\right) \\
&= A \cap (\{Jf > 0\} \cap \mathbb{R}^n) \\
&= A \cap \{Jf > 0\} \\
&= A.
\end{aligned}$$

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

Proof of Claim # 1: Let $g_k = \sum_{i,j=1}^{\infty} \chi_{f(F_j^i)}$ so that $g_k(y)$ is the number of the sets F_j^i , such that $F_j^i \cap f^{-1}\{y\} \neq \phi$. Then, by proof of Lemma 7.0.2, $g_k(y) \longrightarrow \mathcal{H}^0(A \cap f^{-1}\{y\})$ as $k \to \infty$, and

$$\lim_{k\to\infty} \int_{\mathbb{R}^m} g_k(y)\, d\mathcal{H}^n(y) = \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y).$$

Hence, using Theorem 3.0.9 we get

$$\begin{aligned} \lim_{k\to\infty} \int_{\mathbb{R}^m} \sum_{i,j=1}^{\infty} \chi_{f(F_j^i)}\, d\mathcal{H}^n(y) &= \lim_{k\to\infty} \sum_{i,j=1}^{\infty} \int_{\mathbb{R}^m} \chi_{f(F_j^i)}\, d\mathcal{H}^n(y) \\ &= \lim_{k\to\infty} \sum_{i,j=1}^{\infty} \mathcal{H}^n\Big(f\left(F_j^i\right)\Big). \end{aligned}$$

Note that

$$f\restriction_{E_j}\left(F_j^i\right) = f\left(E_j \cap F_j^i\right) = f\left(F_j^i\right).$$

Then,

$$\begin{aligned} \mathcal{H}^n\left(f\left(F_j^i\right)\right) &= \mathcal{H}^n\Big(\left(f\restriction_{E_j} \circ T_j^{-1}\right) \circ \left(T_j\left(F_j^i\right)\right)\Big) \\ &\leq \left(Lip\left(f\restriction_{E_j} \circ T_j^{-1}\right)\right)^n \mathcal{H}^n\left(T_j\left(F_j^i\right)\right). \end{aligned} \tag{7.0.26}$$

Where T_j is as in Lemma 7.0.12. Using lemma 7.0.12 we get

$$\mathcal{H}^n\left(f\left(F_j^i\right)\right) \leq t^n \mathcal{H}^n\left(T_j\left(F_j^i\right)\right).$$

But since $T_j : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ then, $\mathcal{H}^n\left(T_j\left(F_j^i\right)\right) = \mathcal{L}^n\left(T_j\left(F_j^i\right)\right)$. Hence, we conclude that

$$\mathcal{H}^n\left(f\left(F_j^i\right)\right) \leq t^n \mathcal{L}^n\left(T_j\left(F_j^i\right)\right). \tag{7.0.27}$$

Also by Lemma 7.0.2 we have,

$$\begin{aligned} \mathcal{L}^n\left(T_j\left(F_j^i\right)\right) &= \mathcal{H}^n\left(T_j\left(F_j^i\right)\right) \\ &= \mathcal{H}^n\left(T_j \circ \left(f\restriction_{E_j}\right)^{-1} \circ f\left(F_j^i\right)\right) \\ &\leq \left(Lip\left(T_j \circ \left(f\restriction_{E_j}\right)^{-1}\right)^n \circ \mathcal{H}^n\left(f\left(F_j^i\right)\right)\right) \\ &\leq t^n \mathcal{H}^n\left(f\left(F_j^i\right)\right). \end{aligned} \tag{7.0.28}$$

Thus, using 7.0.27 , 7.0.28 and 2.0.2 , and the fact that by Lemma 7.0.3 we have $t^{-n}|det\, T_j| \leq$

$Jf\restriction_{E_j}\leq t^n\,|det\,T_j|$ we get,

$$\begin{aligned}
t^{-2n}\mathcal{H}^n\left(f\left(F_j^i\right)\right) &\leq t^{-n}\mathcal{L}^n\left(T_j\left(F_j^i\right)\right)\\
&= t^{-n}|det\,T_j|\mathcal{L}^n\left(F_j^i\right)\\
&= t^{-n}|det\,T_j|\int_{F_j^i} d\mathcal{L}^n\\
&= \int_{F_j^i} t^{-n}|det\,T_j|\,d\mathcal{L}^n\\
&\leq \int_{F_j^i} Jf\restriction_{E_j}\,d\mathcal{L}^n\\
&= \int_{F_j^i} Jf\,d\mathcal{L}^n\\
&\leq t^n|det\,T_j|\mathcal{L}^n\left(F_j^i\right)\\
&= t^n\mathcal{L}^n\left(T_j\left(F_j^i\right)\right)\\
&\leq t^{2n}\mathcal{H}^n\left(f\left(F_j^i\right)\right).
\end{aligned}$$

Now summing on i and j we get

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

Let $k\to\infty$ and recall **Claim # 1** to get

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

Finally, send $t\to1^+$ to get the equality

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

and we are done.

Case 2: $A\subset\{Jf\left(x\right)=0\}$. Then $\int_A Jf\left(x\right)d\mathcal{L}^n=0$. We will show that $\int_{\mathbb{R}^m}\mathcal{H}^0\left(A\cap f^{-1}\{y\}\right)\,d\mathcal{H}^n=0$. Fix $\epsilon>0$. Let $f=p\circ g$, where

$g:\mathbb{R}^n\longrightarrow\mathbb{R}^n\times\mathbb{R}^m$

$x\longrightarrow\left(f\left(x\right),\epsilon x\right)$ for $x\in\mathbb{R}^n$.

And,

$p:\mathbb{R}^m\times\mathbb{R}^n\longrightarrow\mathbb{R}^m$

$(y,z)\longrightarrow y$ for $y\in\mathbb{R}^m\,,z\in\mathbb{R}^n$.

Then, $p\circ g:\mathbb{R}^n\longrightarrow\mathbb{R}^m$

$x\longrightarrow p\circ g\left(x\right)=p\left(f\left(x\right),\epsilon x\right)=f\left(x\right).$

Claim # 2: There exists a constant C such that $0<Jg\left(x\right)\leq C\epsilon$; for $x\in A$.

Proof of Claim # 2: Write $g=\left(f^1,\cdots,f^m,\epsilon x_1,\cdots,\epsilon x_n\right)$. Then,

$Dg(x) = \begin{pmatrix} Df(x) \\ \epsilon I \end{pmatrix}_{(n+m)\times n}$.

Since $(Jf(x))^2$ is the sum of the squares of $(n \times n)$ subdeterminants of $Df(x)$ according to the Binet-Cauchy formula (see Theorem 6.0.8), then $(Jg(x))^2$ is the sum of the squares of $(n \times n)$ subdeterminants of $Dg(x)$. Let us show that $Jg(x) \geq \epsilon^{2n} > 0$, to see this let
$Dg(x) = \begin{pmatrix} (Df(x))_{m\times n} & \\ & (\epsilon I)_{n\times n} \end{pmatrix}_{(n+m)\times n}$.

Then, $det(\epsilon I) = \epsilon^n$, which implies that $det^2(\epsilon I) = \epsilon^{2n}$. Hence, $(Jg(x))^2 \geq \epsilon^{2n} > 0$. Furthermore, since $|Df| \leq Lip\, f < \infty$, and we may use the Binet-Cauchy formula to compute the following equation $(Jg(x))^2 = (Jf(x))^2 +$ { sum of squares of terms each involving at least one ϵ} $\leq C\epsilon^2$; for each $x \in A$. In order to prove this inequality let

$Df(x) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n-1,2} & \cdots & a_{n-1,n} \end{pmatrix}$. Then, $Dg(x) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ \epsilon & 0 & \cdots & 0 \end{pmatrix}$,
and $|Jg(x)| = \epsilon\, |Df(x)| = C\epsilon$. Then

$$(Jg(x))^2 \leq (c_1\epsilon + c_2\epsilon + \cdots + c_n\epsilon)^2 = \epsilon^2 (c_1 + c_2 + \cdots + c_n)^2 = \epsilon^2 C.$$

Since $p : \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a projection, we can compute using **Case 1**,

$$\begin{aligned} \mathcal{H}^n(f(A)) &= \mathcal{H}^n(p \circ g(A)) \\ &= \mathcal{H}^n(p(g(A))) \\ &\leq (Lip\,p)^n\, \mathcal{H}^n(g(A)). \end{aligned} \tag{7.0.29}$$

Notice that $Lip\,p \leq 1$ thus we get,

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

Let $\epsilon \to 0$, to get $\mathcal{H}^n(f(A)) = 0$. Since the support of $\mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) \subset f(A)$ then,

$$\begin{aligned} \int_{\mathbb{R}^n} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y) &= \int_{spt\, \mathcal{H}^0(A\cap f^{-1}\{y\})} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y) \\ &\leq \int_{f(A)} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y). \end{aligned}$$

But $\mathcal{H}^n(f(A)) = 0$, this implies that $\int_{f(A)} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y) = 0$. This concludes **Case 2**.

Now for the general case let $A = A_1 \cup A_2$ where, $A_1 \subset \{Jf > 0\}$, $A_1 \subset \{Jf = 0\}$. Here we can apply both cases to get, $\int_A Jf(x)\, d\mathcal{L}^n = \int_{\mathbb{R}^m} \mathcal{H}^0\left(A \cap f^{-1}\{y\}\right) d\mathcal{H}^n(y)$. Hence,

$$\int_{A \cap \{Jf>0\}} Jf(x)\, d\mathcal{L}^n + \int_{A \cap \{Jf=0\}} Jf(x)\, d\mathcal{L}^n$$

where the second part of the summand is equal to zero. Thus,

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

□

Chapter 8

Change of Variables formula for $\mathcal{H}^n$

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a Lipschitz function $(n \leq m)$, then for each $\mathcal{L}^n$ - summable function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ we have

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

Remark 8.0.1. Note that using the area formula (see theorem 7.0.4) we notice that $f^{-1}(y)$ is at most countable for $\mathcal{H}^n$ a.e $y \in \mathbb{R}^m$. To see this, we have for all $l \in \mathbb{N}$

$$\begin{aligned} \int_{\mathbb{R}^m} \mathcal{H}^0\left(B(0,l) \cap f^{-1}(y)\right) d\mathcal{H}^n(y) &= \int_{B(0,l)} Jf\, d\mathcal{L}^n \\ &\leq |Lip\ f|^n\ \mathcal{L}^n(B(0,l)) \\ &\leq (Lip\ f)^n\ \alpha_n\ l^n \\ &< \infty. \end{aligned}$$

Since the integral over f is finite hence f is finite $\mathcal{H}^n$ *a.e* ,which implies that

$$\mathcal{H}^0\left(B(0,l) \cap f^{-1}(y)\right) < \infty\ \mathcal{H}^n\, a.e.$$

Notice that $B(0,l) \cap f^{-1}(y)$ is a finite set except on E_l where $\mathcal{H}^n(E_l) = 0$. Let $E = \bigcup_{l=1}^{\infty} E_l$ then

$$\mathcal{H}^n(E) \leq \sum_{l=1}^{\infty} \mathcal{H}^n(E_l) = 0.$$

Let $y \in E^c$ then $y \in \bigcap_{l=1}^{\infty} E_l^c$ which implies that $y \in E_l^c$, $\forall\ l$. Hence, $B(0,l) \cap f^{-1}(y)$ is finite $\forall\, l$. On the other hand,

$$\begin{aligned} f^{-1}(y) &= \mathbb{R}^n \cap f^{-1}(y) \\ &= \left(\bigcup_{l=1}^{\infty} B(0,l)\right) \cap f^{-1}(y) \\ &= \bigcup_{l=1}^{\infty} (B(0,l)) \cap f^{-1}(y) \end{aligned}$$

which is a countable union of finite sets, hence a countable set, thus $f^{-1}(y)$ is at most countable for $\mathcal{H}^n$ a.e $y \in \mathbb{R}^m$.

Proof. 2 cases are to be considered for this proof.

Case 1: If $g \geq 0$ then there exist $\mathcal{L}^n$ measurable sets $\{A_k\}_{k=1}^{\infty}$ in $\mathbb{R}^n$ such that $g = \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}$.
Then, by the Monotone convergence theorem (see theorem 3.0.8) and by the area fomula (see theorem 7.0.4 we have

$$\begin{aligned} \int_{\mathbb{R}^n} g\; Jf\, d\mathcal{L}^n &= \int_{\mathbb{R}^n} \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}\; Jf\, d\mathcal{L}^n \\ &= \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^n} \chi_{A_k}\; Jf\, d\mathcal{L}^n \\ &= \sum_{k=1}^{\infty} \frac{1}{k} \int_{A_k} Jf\, d\mathcal{L}^n \\ &= \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^m} \mathcal{H}^0\left(A_k \cap f^{-1}\{y\}\right)\, d\mathcal{H}^n(y). \end{aligned} \tag{8.0.1}$$

Moreover,

$$\sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^m} \mathcal{H}^0\left(A_k \cap f^{-1}\{y\}\right)\, d\mathcal{H}^n(y) = \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^m} \sum_{x\in f^{-1}\{y\}} \chi_{A_k}(x)\, d\mathcal{H}^n(y) \tag{8.0.2}$$

because, $\mathcal{H}^0\left(A_k \cap f^{-1}\{y\}\right) = \sum_{x\in f^{-1}\{y\}} \chi_{A_k}(x)$. Replacing 8.0.2 in 8.0.1 and interchanging the sum since our functions are positive we get

$$\begin{aligned} \int_{\mathbb{R}^n} g\; Jf\, d\mathcal{L}^n &= \sum_{k=1}^{\infty} \frac{1}{k} \int_{\mathbb{R}^m} \sum_{x\in f^{-1}\{y\}} \chi_{A_k}(x)\, d\mathcal{H}^n(y) \\ &= \int_{\mathbb{R}^m} \sum_{x\in f^{-1}\{y\}} \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x)\, d\mathcal{H}^n(y) \\ &= \int_{\mathbb{R}^m} \sum_{x\in f^{-1}\{y\}} g(x)\, d\mathcal{H}^n(y). \end{aligned}$$

Case 2: If g is any $\mathcal{L}^n$ - summable function, then g can be written as the sum of two positive functions, let $g = g^+ - g^-$. Now applying **case 1** on g^+ and g^- we get

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

and

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

hence,

$$\begin{aligned} \int_{\mathbb{R}^n} g(x) \ Jf(x) \ d\mathcal{L}^n &= \int_{\mathbb{R}^n} \left(g^+ - g^-\right) \ Jf \, d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} g^+(x) \ Jf(x) \ d\mathcal{L}^n - \int_{\mathbb{R}^n} g^-(x) \ Jf(x) \ d\mathcal{L}^n, \end{aligned}$$

the last equality comes from the fact that g^+ and g^- are $\mathcal{L}^n$ summable on $\mathbb{R}^n$. And hence,

$$\begin{aligned} \int_{\mathbb{R}^n} g^+(x) \ Jf(x) \ d\mathcal{L}^n - \int_{\mathbb{R}^n} g^-(x) \ Jf(x) \ d\mathcal{L}^n &= \int_{\mathbb{R}^m} \sum_{x \in f^{-1}\{y\}} g^+(x) \ d\mathcal{H}^n(y) - \int_{\mathbb{R}^m} \sum_{x \in f^{-1}\{y\}} g^-(x) \ d\mathcal{H}^n(\\ &= \int_{\mathbb{R}^m} \sum_{x \in f^{-1}\{y\}} \left(g^+ - g^-\right)(x) \ d\mathcal{H}^n(y) \\ &= \int_{\mathbb{R}^m} \sum_{x \in f^{-1}\{y\}} g(x) \ d\mathcal{H}^n(y). \end{aligned}$$

□

Chapter 9

Applications of the Area Formula

A- Length of a curve : $(n = 1; m \geq 1))$.
Consider any injective Lipschitz function $f : \mathbb{R} \longrightarrow \mathbb{R}^m$, and consider the curve $C = f([a, b]) \subset \mathbb{R}^m$, where $-\infty < a < b < \infty$. Using the area formula, we show that the length of the curve C is $\mathcal{H}^1(C) = \int_a^b Jf \, d\mathcal{L}^n$, where $Jf = |Df|$.

Proof. By the area formula, we have

$$\begin{aligned} \int_a^b Jf \, d\mathcal{L}^n &= \int_{\mathbb{R}^m} \mathcal{H}^0 \left([a, b] \cap f^{-1}(\{y\})\right) d\mathcal{H}^1(y) \\ &= \int_{\mathbb{R}^m \cap f([a,b])} \mathcal{H}^0 \left([a, b] \cap f^{-1}(\{y\})\right) d\mathcal{H}^1(y) + \int_{\mathbb{R}^m \setminus f([a,b])} \mathcal{H}^0 \left([a, b] \cap f^{-1}(\{y\})\right) d\mathcal{H}^1(y) \end{aligned} \tag{9.0.1}$$

Notice that the second part of the summand in (9.0.1) is zero, since for $y \in \mathbb{R}^m \setminus f([a, b])$, $[a, b] \cap f^{-1}(\{y\}) = \emptyset$, and thus $\mathcal{H}^0([a, b] \cap f^{-1}(\{y\})) = 0$. As for the first part of the summand, we recall that f is injective, and thus for $y \in f([a, b])$, there exists a unique $x \in [a, b]$ such that $f(x) = y$. Hence, in this case, we get that $\mathcal{H}^0([a, b] \cap f^{-1}(\{y\})) = 1$. Therefore, plugging in equation (9.0.1), we get

$$\int_a^b Jf \, d\mathcal{L}^n = \int_{\mathbb{R}^m \cap f([a,b])} 1 \, d\mathcal{H}^1(y) = \mathcal{H}^1(f([a, b])) = \mathcal{H}^1(C)$$

□

B- Surface area of a graph:$(n \geq 1; m = n + 1)$.
Let $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ be any Lipschitz function. For $U \subset \mathbb{R}^n$ open set define the graph of g over U to be $G = \{(x, g(x)), x \in U\}$. Then,

$$\begin{aligned} \mathcal{H}^{n+1}(G) &:= \text{Surface area of G} \\ &= \int_U \left(Jf^2\right)^{\frac{1}{2}} d\mathcal{L}^n. \end{aligned}$$

Proof. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$, be defined as $f(x) = (x, g(x))$. Notice that f is Lipschitz since

$$\begin{aligned} |f(x) - f(y)| &= |(x, g(x)) - (y, g(y))| \\ &= |(x - y, g(x) - g(y))| \\ &\leq (1 + Lip\, g)\, |x - y|. \end{aligned}$$

Moreover, note that $Df = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \\ 0 & 0 & \cdots & 1 \\ \frac{dg}{dx_1} & \frac{dg}{dx_2} & \cdots & \frac{dg}{dx_n} \end{pmatrix}_{(n+1)\times n}$

Now, we need to prove that $(Jf)^2 = 1 + |Dg|^2$. To simplify our calculations , let us take a small example : Suppose $g : \mathbb{R}^3 \longrightarrow \mathbb{R}$ and define $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$, then $Df = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{dg}{dx_1} & \frac{dg}{dx_2} & \frac{dg}{dx_3} \end{pmatrix}_{4\times 3}$;

and by definition $(JF)^2 =$ sum of squares of 3×3 subdeterminants so that :

$$(JF)^2 = \left(det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)^2 + \left(det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{dg}{dx_1} & \frac{dg}{dx_2} & \frac{dg}{dx_3} \end{pmatrix}\right)^2 + \left(det \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{dg}{dx_1} & \frac{dg}{dx_2} & \frac{dg}{dx_3} \end{pmatrix}\right)^2 + \left(det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \frac{dg}{dx_1} & \frac{dg}{dx_2} & \frac{dg}{dx_3} \end{pmatrix}\right)^2$$

$$\begin{aligned} &= 1^2 + \left(\frac{dg}{dx_3}\right)^2 + \left(\frac{dg}{dx_1}\right)^2 + \left(\frac{dg}{dx_2}\right)^2 \\ &= 1 + |Dg|^2. \end{aligned}$$

Now for the general case; if we have an $(n+1) \times n$ matrix, then by taking the sum of squares of all $n \times n$ subdeterminants we will end up by getting $(Jf)^2 = 1 + \left(\frac{dg}{dx_1}\right)^2 + \cdots + \left(\frac{dg}{dx_n}\right)^2$, which is nothing but $1 + |Dg|^2$.

Using the area formula we have that $\int_U Jf\, d\mathcal{L}^n = \int_{\mathbb{R}^{n+1}} \mathcal{H}^0\left(U \cap f^{-1}(y)\right) d\mathcal{H}^n(y)$.

$$\begin{aligned} \text{but } \int_U Jf\, d\mathcal{L}^n &= \int_U \left(1 + |Dg|^2\right)^{\frac{1}{2}} d\mathcal{L}^n \\ &= \int_{\mathbb{R}^{n+1} \cap (U \times g(U))} \mathcal{H}^0\left(U \cap f^{-1}(y)\right) d\mathcal{H}^n(y) + \int_{\mathbb{R}^{n+1} \setminus (U \times g(U))} \mathcal{H}^0\left(U \cap f^{-1}(y)\right) d\mathcal{H}^n(y) \qquad (10.2) \end{aligned}$$

Notice that the second part of the summand in (9.0.2) is zero, since if there exists $x \in U \cap f^{-1}(y)$ then , $f(x) = y = (x, g(x))$ and $x \in U$ implies that $y \in U \times g(U)$ which is a contradiction. As for the first summand, notice that f is one-to-one:
Suppose $f(x_1) = f(x_2)$ then $(x_1, g(x_1)) = (x_2, g(x_2))$ this implies $x_1 = x_2$ and $g(x_1) = g(x_2)$. Hence f is one-to-one which implies that $\mathcal{H}^0\left(U \cap f^{-1}(y)\right) = 1$. Thus,

$$\begin{aligned}\int_{\mathbb{R}^{n+1}\cap(U\times g(U))} \mathcal{H}^0\left(U \cap f^{-1}(y)\right) d\mathcal{H}^n(y) &= \int_{\mathbb{R}^{n+1}\cap(U\times g(U))} 1\, d\mathcal{H}^n(y) \\ &= \int_{U\times g(U)} d\mathcal{H}^n(y) \qquad (9.0.3)\end{aligned}$$

But $U \times g(U)$ is $f(U)$ hence, replacing 9.0.3 in 9.0.2, we get

$$\int_U Jf(x)\, d\mathcal{L}^n = \int_{U\times g(U)} d\mathcal{H}^n(y) = \int_{f(U)} d\mathcal{H}^n(y) = \mathcal{H}^n(f(U)) = \mathcal{H}^n(G).$$

□

C- surface area of a Parametric hypersurface $(n \geq 1, m = n+1)$.
Consider any one-to-one Lipschitz function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$. Let $U \subseteq \mathbb{R}^n$ be an open set and $S = f(U) \subset \mathbb{R}^{n+1}$. Then,

$$\mathcal{H}^n(S) = \int_U [(Jf)^2]^{\frac{1}{2}}\, d\mathcal{L}^n \text{ ; where } (Jf)^2 = \text{ sum of square of } n \times n$$

$$\text{subdeterminants of the } (n+1)\times n \text{ matrix} = \sum_{K=1}^{n+1}[\frac{\partial\left(f^1, \cdots, f^{k-1}, f^{k+1}, \cdots, f^{n+1}\right)}{\partial(x_1, \cdots, x_n)}]^2.$$

Proof. Write $f = (f^1, \cdots, f^{n+1})$ where each $f^i : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a Lipschitz function.Let us calculate $(Jf)^2$.

Note that $Df = \begin{pmatrix} \frac{\partial f^1}{\partial x_1} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f^{n+1}}{\partial x_1} & \cdots & \frac{\partial f^{n+1}}{\partial x_n} \end{pmatrix}_{(n+1)\times n}$

First, let us take a small example on how to derive the formula of the Jacobian. For the simplicity of calculations, suppose n=2, then $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$;

$$Df = \begin{pmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2} \\ \frac{\partial f^3}{\partial x_1} & \frac{\partial f^3}{\partial x_3} \end{pmatrix}_{3\times 2}$$

then $(Jf)^2 = \left(det\begin{pmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2}\end{pmatrix}\right)^2 + \left(det\begin{pmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} \\ \frac{\partial f^3}{\partial x_1} & \frac{\partial f^3}{\partial x_2}\end{pmatrix}\right)^2 + \left(det\begin{pmatrix} \frac{\partial f^2}{\partial x_1} & \frac{\partial f^3}{\partial x_1} \\ \frac{\partial f^2}{\partial x_2} & \frac{\partial f^3}{\partial x_2}\end{pmatrix}\right)^2$

which is equivalent to $\left[\frac{\partial(f^1, f^2)}{\partial(x_1, x_2)}\right]^2 + \left[\frac{\partial(f^1, f^3)}{\partial(x_1, x_2)}\right]^2 + \left[\frac{\partial(f^2, f^3)}{\partial(x_1, x_2)}\right]^2$.

For the general case we have that for the $(n+1)\times n$ matrix , the formula of the Jacobian means,that for each K , the $(n\times n)$ subdeterminant is the determinant of the partial derivatives of f ; that is the partial derivatives of $(f^1,\cdots,f^{n+1})$ with respect to $(x_1,\cdots,x_n)$ except for the k^{th} one.

Now coming back to the application of the area formula we get that

$$\begin{aligned}\int_U Jf\,d\mathcal{L}^n &= \int_{\mathbb{R}^{n+1}} \mathcal{H}^0\left(U\cap f^{-1}(y)\right)\,d\mathcal{H}^n(y)\\ &= \int_{\mathbb{R}^{n+1}\cap f(U)} \mathcal{H}^0\left(U\cap f^{-1}(y)\right)\,d\mathcal{H}^n(y)+\int_{\mathbb{R}^{n+1}\setminus f(U)} \mathcal{H}^0\left(U\cap f^{-1}(y)\right)\,d\mathcal{H}^n(y) \quad (9.0.4)\end{aligned}$$

The second part of the summand in (9.0.4) is zero, since if $x\in U\cap f^{-1}(y)$ then $x\in U$ and $f(x)=y$, thus $y\in f(U)$ which is a contradiction. And since f is one to one then

$$\mathcal{H}^0\left(U\cap f^{-1}(y)\right)\,dH^n(y)=1.$$

Hence the first part of the summand implies that

$$\begin{aligned}\int_{\mathbb{R}^{n+1}\cap f(U)} d\mathcal{H}^n(y) &= \int_{f(U)} d\mathcal{H}^n(y)\\ &= \mathcal{H}^n\left(f(U)\right)\\ &= \mathcal{H}^n\left(S\right).\end{aligned}$$

□

Bibliography

[1] Lawrence C. Evans and Ronald F. Gariepy, *Measure Theory and Fine Properties of Functions*, (Studies in Advanced Mathematics, CRC Press, 1992).

[2] Michael Spivak , *Calculus on Manifolds*, (Addison-Wesley Publishing Company , The advanced Book Program , 1965).

[3] Elias M. Stein and Rami Shakarchi , *Real Analysis Measure Theory , Integration and Hilbert Spaces*, (Pinceton University Press , Princeton and Oxford , 2005).

[4] Gerald B. Folland , *Real Analysis, Modern Techniques and Their Applications*, (John Wiley and sons INC. , 1999).