Fine Properties of measurable functions

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Abhishek Adimurthi

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April , 2020

Supervisor: Dr. Shyam Sundar Ghoshal

Certificate

This is to certify that this dissertation entitled : Fine Properties of measurable functions, towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune, represents study/work carried out by Abhishek Adimurthi at Tata Institute for Fundamental Research, Center for Applicable Mathematics, Bengaluru, under the supervision of Dr. Shyam Sundar Ghoshal, Reader, during the academic year 2019-2020.

Dr. Shyam Sundar Ghoshal.

Committee :
Dr. Shyam Sundar Ghoshal

Dr. Debdip Ganguly
Debdip Ganguly

Declaration

I hereby declare that the matter embodied in the report entitled : Fine Properties of measurable functions, are the results studied by me at TIFR CAM, Bengaluru, under the supervision of Dr. Shyam Sundar Ghoshal and the same has not been submitted elsewhere for any other degree.

Abhishek

Abhishek Adimurthi

Acknowledgement

I would like to thank my professors at IISER Pune, for their time and their valuable teaching, from whom I got inspired for the little bit of knowledge of mathematics, I now know of.
Prof. Debdip Ganguly and Prof. Mousomi Bhakta were some of the professors who inspired me to take up measure theory and PDE as my topic of reasearch interest. Prof. A Raghuram inspired me by his vast mathematical knowledge and his style in teaching. I was also inspired by Prof. Chandrasheel Bhagwat for his preciseness in teaching and his spread-out interest in mathematics. I would like to thank Prof. K T Joseph, Prof. G D V Gowda and Prof. K Sandeep, for helping me increase my knowledge in PDE and advanced functional analysis and inspiring me to be open to any kind of new proofs and the precision towards it.
Pages and type-ups are definitely not suffecient to express my gratitude for the professors at IISER and TIFR, but I would like to thank them all for providing their valuable time to expand my knowledge.
I would like to express my gratitude to Prof. Shyam Sundar Ghoshal, my project supervisor for the thesis, for his valuable time and patience to teach me and guide me for my MS Thesis. I was inspired by his frankness and his precision for the proofs.
I would like to thank my friends, especially my parents and my brother for tolerating me and my stupid, time consuming doubts, ideas and arguments.
I would like to thank TIFR -CAM in all aspects, for providing me the oppurtunity to do my thesis at TIFR CAM.
Last, but not the least, I would like to thank all those who helped me do better mathematics or in life, in general.

Introduction

This document compiles some theorems and proofs primarily related to measure-theory and geometry in view of a part of my reading project done at TIFR CAM, in the context of my MS Thesis at IISER Pune.
With the assumptions of the basic knowledge on Measure Theory, Functional Analysis, and Linear Algebra, this document proves the extension of measures from a variety of family of sets, Taylor's formula, the Change of variables among the spaces of the same dimension, some properties of functions which are absolutely continuous and the Integration by parts. It also gives an insight via a compilation of some theorems and proof, of, the Fubini and Tonelli theorem(without the use of the monotone class lemma), the Radon-Nikodym theorem and the Radon Nikodym derivative, some covering theorems and some of the important properties of Lipschitz functions, namely the Rademacher theorem. It also has information on the dimension of fractals (going with the name of Hausdorff dimension), the Isodiameteric inequality, which is further used to prove the change of variables formula among spaces of a different dimensions, which goes by the name of the Area and the Co-Area Formula.
Most of the materials in this document are taken from the references mentioned at the end of this document.

Contents

Dedicated to my very tolerant and lovely, Mom and Dad,
and my Brother.

Some Notations and Conventions

$\mathbb{R}^n$	The n dimensional euclidean space.
$dx \equiv dx_1.dx_2 \ldots dx_n$	The n-dimensional lebesgue measure on $\mathbb{R}^n$
$\mathbb{C}^n$	The n-dimensional complex space.
$\hat{B}$	The ball with the center as that of B, but with the radius 5 times of B
$GL_n(\mathbb{R})$	The group of $n \times n$ invertible matrices with real entries.
$\overline{A}$	The topological closure of the set A.
A°	The topological interior of the set A.
∂A	The boundary of the set $A := \overline{A} - A^\circ$
m	Measure defined in the definition
$(m\|A)$	Measure restricted to the set A as in definition
$D_u v$	Density of v with respect to u
$\mathcal{L}^n$	Lebesgue measure on $\mathbb{R}^n$
$\alpha(n)$	Lebesgue measure of a unit ball in $\mathbb{R}^n$
$Card(K)$	Cardinality of the set K
MCT	Monotone Convergence Theorem
DCT	Dominated Convergence Theorem
$\mathcal{X}_A$	Charecteristic function with respect to the set A
$\mathcal{H}^n$	The n dimensional Hausdorff measure.
$diam(B)$	The diameter of B
$S_a(A)$	Steiner symmetrisation of A w.r.t the plane P_a
$\int^* f$	$\inf\{\int \phi \;;\; \phi$ is simple and $f \leq \phi\}$
$[[L]]$	Jacobian of L
$C_c(X)$	Real valued continuous functions with compact support on X
$C_c^k(X)$	Real valued k times differentiable functions with compact support on X
a.e	almost everywhere

Chapter 1

Pre-Requisites

1.1 Linear Algebra

- V is a Vector Space over a field $\mathbb{F}$, generally considered over $\mathbb{R}$ or $\mathbb{C}$,
  if $\forall x, y \in V$ and $\forall \alpha, \beta \in \mathbb{F}$,

  $$\text{Denote } \alpha x + \beta y \in V \text{ as the joint binary operation on } V$$

- Basis of a vector space V is defined as a collection of elements of the vector space such that the collection of elements are linearly independent and spans the whole V by finite linear combinations.

- V is a finite dimensional vector space if there exists a basis for V which is finite by cardinality.

- If V and W are two vector spaces, T is a linear transformation from V to W if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$, with the usage of notations as in previous points.

- A subspace S is a subset of a vector space V and is a vector space itself, that is, $S \subset V$ and for all $x, y \in S$, $\alpha, \beta \in \mathbb{F}$, $\alpha x + \beta y \in S$.

- $ker(T) := \{x \in V \mid T(x) = 0\}$ and $Range(T) := \{y \in W \mid \exists x \in V, T(x) = y\}$ are examples for subspaces of V and W respectively.

- If V is finite dimensional vector space and given any two basis of it, one can transform from one basis to another in matrix representation, by an invertible matrix called change of basis matrix.

- Rank-Nullity theorem : For a linear transform defined earlier on a finite dimensional vector space V, we have

$$\dim(V) = \dim(ker(T)) + \dim(Im(T))$$

- For a subspace W of V, W is said to be invariant under T if $TW \subset W$.
- A complex (Real) vector space H is called an inner product space if $\forall x, y \in H$, there is an associated complex number (real number), denoted by $< x, y >$ called the inner product with the properties:
  - $< x, y >= \overline{< y, x >}$
  - $< x + y, z >=< x, z > + < y, z >$
  - $< \alpha x, y >= \alpha < x, y >$ for all $x, y \in H, \alpha \in \mathbb{C}$
  - $< x, x > \geq 0$ for all $x \in H$
  - $< x, x >= 0 \iff x = 0$

  One can define norm on H by $||x||^2 =< x, x >$ and hence, one can define distance on H by $d(x, y) = ||x - y||$
- Cauchy Schwartz inequality: For a inner product space defined in the earlier point,

$$| < x, y > | \leq ||x|| \; ||y|| \text{ ; for all } x, y \in H$$

- H is a Hilbert space if it has an inner product structure and is complete by the convergence of every cauchy sequence in H with respect to the metric defined above.
- If H is a Hilbert space over $\mathbb{R}$, then H is called real Hilbert space.
- Every linear map T over finite dimensional space V, spanned by the basis $\{e_1, e_2, \ldots, e_n\}$, to a finite dimesional space W, spanned by the basis $\{f_1, f_2, \ldots f_m\}$, can be represented by a $m \times n$ matrix which is given by

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m \times n}$$

  where a_{ij} is defined by

$$T(e_i) = \sum_{j=1}^{m} a_{ji} f_j \text{ for all } 1 \leq i \leq n$$

- Let $GL_n(\mathbb{R}^n)$ be the set of all invertible matrices over $\mathbb{R}$. Any $T \in GL_n(\mathbb{R}^n)$ can be written as product of finitely many elementary transformations of the types:
  - I : $T_1(x_1, x_2, \ldots, x_i, \ldots x_n) = (x_1, x_2, \ldots, cx_i, \ldots x_n)$; $c \neq 0$
  - II : $T_2(x_1, x_2, \ldots, x_i, \ldots x_n) = (x_1, x_2, \ldots, x_i + cx_k, \ldots x_n)$, $k \neq i$
  - III : $T_3(x_1, x_2, \ldots, x_i, \ldots, x_j \ldots x_n) = (x_1, x_2, \ldots, x_j, \ldots, x_i \ldots x_n)$
- Let S_n denote all bijections from $\{1, 2, \ldots n\}$ to $\{1, 2, \ldots n\}$. S_n forms a group under composition. Any element of S_n, called as permutation, can be decomposed as product of transpositions, where a transposition is defined as an element of S_n which just interchanges two elements of $\{1, 2, \ldots n\}$, keeping the other elements fixed. For a $\sigma \in S_n$, we define $sgn(\sigma) = (-1)^m$; where m is the number of transpositions in the decomposition of σ. Here,for an element of S_n, the sgn is invariant of representation by composition of transpositions.
  An important property of sgn is for any two elements σ and τ in S_n, we have $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$
- Determinant of a matrix A of order $n \times n$: is a multinlinear map D from $\mathbb{C}^n \times \mathbb{C}^n \times \ldots \mathbb{C}^n{}_{\text{n times}}$ to $\mathbb{C}$ with the property : for all $R_i, L_i \in \mathbb{C}^n$, $1 \leq i \leq n$ and for all $\alpha, \beta \in \mathbb{C}$
  - Multilinear : For all $1 \leq i \leq n$,

$$D(R_1, R_2, \ldots, \alpha R_i + \beta L_i, \ldots, R_n)$$

$$||$$

$$\alpha D(R_1, R_2, \ldots, R_i, \ldots, R_n) + \beta D(R_1, R_2, \ldots, L_i, \ldots, R_n)$$

  - Alternating : For all $1 \leq i, j \leq n$,

$$D(R_1, R_2, \ldots, R_i, \ldots, R_j, \ldots, R_n) = (-1)D(R_1, R_2, \ldots, R_j, \ldots, R_i, \ldots, R_n)$$

  - Define $E_i = (0, \ldots, 0, 1_{\text{ith place}}, 0, \ldots, 0) \in \mathbb{C}^n$, we have

$$D(E_1, E_2, \ldots, E_n) = 1$$

- Some properties of determinant includes :
  - Let A be a $n \times n$ matrix. The determinant of A, denoted by $det(A)$, is defined to be $D(R_1, R_2, \ldots, R_n)$, where R_i is the ith row of A.
  - Let $\sigma \in S_n$, then

$$D(R_{\sigma(1)}, R_{\sigma(2)}, \ldots, R_{\sigma(n)}) = sgn(\sigma)D(R_1, R_2, \ldots, R_n)$$

- Let $A \equiv [a_{ij}]_{n\times n}$, then

$$det(A) = \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)}$$

- $$det(AB) = det(A).det(B)$$

- For A to be invertible,
$$det(A^{-1}) = det^{-1}(A)$$

- If A^T is the transpose of the matrix $A \equiv (a_{ij})$,
given by (a_{ji}), then
$$detA = det(A^T)$$

- For α to be an element in the field,

$$\det(\alpha A) = \alpha^n det(A)$$

- Determinant of diagonal matrix, triangular matrix (upper triangular or lower triangular matrix) is product of elements of the principle diagonal elements of the matrix, that is, if the matrix of order $n \times n$ is given by $(a_{ij})_{n\times n}$, then the determinant of it is $\prod_{i=1}^{n} a_{ii}$
- Invariance of determinant with regard to the basis:
For a linear transformation S and for a basis of the $n-$dimensional vector space V, say $\{e_1, e_2, \ldots, e_n\}$, denote the matrix representation by $[S]$. For $\{f_1, f_2, \ldots, f_n\}$ to be another basis, and $[S']$ to be the representation of S in the new basis, then there exists an invertible matrix, P, such that $[S] = P \circ [S'] \circ P^{-1}$.
And hence, by the properties of determinant, mentioned earlier, we have

$$det([S]) = det([S'])$$

- For a linear transformation S and the basis $\{e_1, e_2, \ldots, e_n\}$, denote the matrix representation by $[S]$. Then the chareceteristic polynomial of S is a polynomial of degree n with the variable λ, given by

$$det\Bigg(\lambda.I_{n\times n} - [S]\Bigg)$$

The roots of the above polynomial are called the eigenvalues of $[S]$ and $det([S])$ is precisely the product of eigenvalues. The trace of $[S]$ denoted by $Tr([S])$ is given by $\sum_{i=1}^{n}[S]_{ii}$ i.e the sum of the diagonal elements of $[S] \equiv \big([S]_{ij}\big)_{n\times n}$, which is precisely equal to the sum of the eigenvalues of $[S]$.

- For an inner product space V, with the inner product defined by $<,>$, we say $a, b \in V$ is orthogonal, denoted by $a \perp b$ if $< a, b >= 0$.
  Two sets A and B are said to be orthogonal, if $a \perp b$ for all $a \in A$, $b \in B$.

- For a linear transformation T from H to H as defined earlier with H to be a finite dimensional Hilbert space, we define the adjoint of T, denoted by T^*, as a linear map which satisfies

$$< Tx, y >=< x, T^*y > \;\; \forall\; x, y \in H$$

- Let $\{V, <,>\}$ and $\{W, (,)\}$ be two finite dimensional Hilbert spaces with the corresponding inner products $<,>, (,)$.
  O is called an orthogonal transformation if for all $v_1, v_2 \in V$, we have

$$(O(v_1), O(v_2)) =< v_1, v_2 >$$

- Let $O : V \to W$ be an orthogonal transformation and let $v_1, v_2 \in V$, then

$$< v_1, v_2 >= (O(v_1), O(v_2)) =< O^*Ov_1, v_2 >$$

- Some properties of the orthogonal transformation :

  - $O^*Ov_1 = v_1$, for all v_1 in V. Hence $O^*O = I_V$, where I_V denote the identity map from V to V.
  - Let W be the range of O. Then there is a $x \in V$ such that $w = O(x)$. Hence,

$$OO^*(w) = OO^*(Ox) = O(O^*O(x)) = O(x) = w$$

    Thus OO^* is precisely the identity map on the range of O.
  - With the orthogonal basis of the vector space fixed, O represents a matrix called the orthogonal matrix.

- Let H be a finite dimensional Hilbert space with the inner product $<,>$. Let the dimension of H be n. Let V, W be finite dimensional subspaces of H such that $k = dim(V) = dim(W)$. Let $T : V \to W$ be an orthogonal transformation. Then, there is an orthogonal transformation $Q : H \to H$ such that $Q = T$ on V.
  In order to prove this, let $\{e_1, e_2, \ldots, e_n\}$ and $\{f_1, f_2, \ldots, f_n\}$ be 2 orthonormal basis for H such that

- $span\{e_1, e_2, \ldots, e_k\} = V$
- $Te_i = f_i$ for $1 \leq i \leq k$.

Define $Q : H \to H$ by $Q(e_i) = f_i$ for all $1 \leq i \leq n$. Then Q is orthogonal and $Q(e_i) = T(e_i)$ for all $1 \leq i \leq k$. Hence, Q is the required orthogonal transformation.

- Spectral theorem : For V a finite dimensional Hilbert space with $<,>$ as inner product, let T be a linear, self adjoint operator i,e $T^* = T$ and $p(\lambda)$ be the charecteristic polynomial of T, with $\lambda_1, \ldots \lambda_k$ having multiplicities $m_1, m_2, \ldots m_k$ be the distinct eigenvalues of T.
  Let
  $$\text{The } \lambda_i \text{ eigen space of T} : V_{\lambda_i} = \{v \in V \mid Tv = \lambda_i v\}$$
  then
  - $dim(V_{\lambda_i}) = m_i, \forall i$
  - $\lambda_i \in \mathbb{R}, \forall i$
  - $V_{\lambda_i} \perp V_{\lambda_j}, i \neq j$
  - $V = V_{\lambda_1} \bigoplus \cdots \bigoplus V_{\lambda_k}$
- Remark to the above point: If A is a real symmetric matrix, the above theorem holds for A and also that A is diagonalizable, that is, there is P orthogonal matrix and D diagonal matrix such that $A = P \circ D \circ P^{-1}$.
- Riesz Representation theorem : Let $\big(V, <,> \big)$ be a finite dimensional Hilbert space and $L : V \to V$ be a linear transformation.
  Then $\exists! w \in V$ s.t $L(v) =< v, w >, \forall v \in V$.
- Define (standard) inner product for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, called the dot product of x and y as
  $$x.y = \sum_1^n x_i y_i \in \mathbb{R}$$

1.2 Calculus involved in Linear Algebra

Theorem 1.1. ***Polar Decomposition :***
For $L : \mathbb{R}^n \to \mathbb{R}^m$, linear transformation, L can be decomposed as the following, in the cases given below :

- $n \leq m$,
$$L = O \circ S$$
where $O : \mathbb{R}^n \to \mathbb{R}^m$; *orthogonal,* $S : \mathbb{R}^n \to \mathbb{R}^n$ *Symmetric*

- $n \geq m$,
$$L = S \circ O^*$$
where $O : \mathbb{R}^m \to \mathbb{R}^n$; *orthogonal,* $S : \mathbb{R}^m \to \mathbb{R}^m$ *Symmetric*

Proof. Consider the case $n \leq m$. Consider $C = L^* \circ L : \mathbb{R}^n \to \mathbb{R}^n$

$$(Cx).y = (L^* \circ Lx).y = Lx.Ly = x.(Cy)$$

Hence C is symmetric.
$$(Cx).x = Lx.Lx \geq 0$$

And hence the eigenvalues of C are non negative. This concludes that C is symmetric and non negative definite. Hence, there are $\lambda_1, \ldots, \lambda_n$ positive and there is an orthonormal basis $\{x_1, \ldots x_n\}$ of $\mathbb{R}^n$ with $Cx_i = \lambda_i x_i$
W.L.O.G assume that $\lambda_i \neq 0$ for $i < s$ and $\lambda_i = 0$ for $s \leq i \leq n$.
For $i < s$, define
$$u_i = +\sqrt{\lambda_i}$$
$$\xi_i = L\left(\frac{x_i}{u_i}\right)$$

Then, for $1 \leq i, j < s$,

$$\xi_i.\xi_j = \frac{1}{u_i u_j} L(x_i).L(x_j) = \frac{1}{u_i u_j} L^* L x_i.x_j = \frac{u_i}{u_j} x_i.x_j = \frac{u_i}{u_j} \delta_{ij}$$

Hence, $\{\xi_1, \ldots, \xi_{s-1}\}$ is an orthonormal set of $\mathbb{R}^m$. Let $\{\xi_s, \ldots, \xi_m\} \subset \mathbb{R}^m$ be such that $\{\xi_1, \ldots, \xi_m\}$ forms the orthonormal basis of $\mathbb{R}^m$.

Declare S on $\{x_j\}$ as
$$Sx_j = u_j x_j$$

and O as
$$Ox_j = \xi_j$$

Then O is orthogonal and S is symmetric. Furthermore, for $1 \leq i < s$,

$$(O \circ S)(x_i) = O\big(u_i x_i\big) = u_i O(x_i) = u_i \frac{Lx_i}{u_i} = Lx_i$$

And, for $i \geq s$,

$$\big(O \circ S\big)(x_i) = O(0) = 0 = Lx_i$$

Hence, $L = O \circ S$ and this completes the proof for the case $n \leq m$.
For the case $n \geq m$, repeat the proof for $L^* : \mathbb{R}^m \to \mathbb{R}^n$ to get the corresponding S and O. □

Definition 1.2. Assume that $L : \mathbb{R}^n \to \mathbb{R}^m$ be linear.

- If $n \leq m$, declare $L = O \circ S$ as above and we define the Jacobian of L to be

$$[[L]] = |detS|$$

- If $n \geq m$, declare $L = S \circ O^*$ as above and we define the Jacobian of L to be

$$[[L]] = |detS|$$

Theorem 1.3.

- *If* $n \leq m$, *then*

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

- *If* $n \geq m$, *then*

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

Proof. For the case $n \leq m$, declare $L = O \circ S$. Then $L^* = S \circ O^*$ and hence

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

Using the property of orthogonal maps, that is $O^* \circ O = I$, we get,

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

The case of $n \geq m$, is similar to the previous one, but consider $L \circ L^*$ and repeat the proof as above. □

Remark 1.4. If L has another representation of S and O, then the previous theorem says that the $[[L]]^2$ is independent of the representations and $[[L]] = [[L^*]]$.

Definition 1.5. Let L be a linear map from $\mathbb{R}^n \to \mathbb{R}^m$,

- If $n \leq m$, then define

$$\Lambda(m,n) := \{\lambda : \{1,\dots,n\} \to \{1,\dots,m\} \mid \lambda \text{ is increasing } \}$$

- For each $\lambda \in \lambda(m,n)$, declare $P_\lambda : \mathbb{R}^m \to \mathbb{R}^n$ as

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

Theorem 1.6. ***Binet - Cauchy Formula :***
Assume $n \leq m$ and L be a linear map from $\mathbb{R}^n \to \mathbb{R}^m$, Then

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

Proof. With respect to the standard basis on $\mathbb{R}^n$ and $\mathbb{R}^m$, write the matrix as

$$L = \big((l_{ij})\big)_{m \times n}$$

$$A = L^* \circ L = \big((a_{ij})\big)_{n \times n}$$

Hence, the $(i,j)^{th}$ element of A turns out to be precisely

$$a_{ij} = \sum_{k=1}^{m} l_{ki} l_{kj}$$

Thus, by the definition of determinant,

$$[[L]]^2 = detA = \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

where $\mathcal{S}_n$ denotes the permutations of $\{1,2,\dots,n\}$. Hence

$$[[L]]^2 = \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} \sum_{k=1}^{m} l_{ki} l_{k\sigma(i)}$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} [l_{1i} l_{1\sigma(i)} + l_{2i} l_{2\sigma(i)} + \cdots + l_{mi} l_{m\sigma(i)}]$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \left\{ \begin{array}{c} [l_{11}l_{1\sigma(1)} + l_{21}l_{2\sigma(1)} + \cdots + l_{m1}l_{m\sigma(1)}] \\ \times \\ [l_{12}l_{1\sigma(2)} + l_{22}l_{2\sigma(2)} + \cdots + l_{m2}l_{m\sigma(2)}] \\ \times \\ \vdots \\ \times \\ [l_{1n}l_{1\sigma(n)} + l_{2n}l_{2\sigma(n)} + \cdots + l_{mn}l_{m\sigma(n)}] \end{array} \right\}$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma)[\sum_{\phi \in \Phi} \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)} + \sum_{\phi \notin \Phi} \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)}]$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \sum_{\phi \in \Phi} \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)} + \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \sum_{\phi \notin \Phi} \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)}$$

$$:= I_1 + I_2$$

where $\Phi := \Big\{$ one - one mappings of $\{1, 2, \ldots, n\}$ into $\{1, 2, \ldots, m\}$ $\Big\}$.
Finite sums can be interchanged and hence,

$$I_2 = \sum_{\phi \notin \Phi} \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)}$$

$$= \sum_{\phi \notin \Phi} \left(\prod_{i=1}^{n} l_{\phi(i),i} \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} l_{\phi(i),\sigma(i)} \right)$$

For $\phi \notin \Phi$, denote the matrix $A(\phi) = \big((a_{ij}(\phi))\big)$, where

$$a_{ij}(\phi) = l_{\phi(i),j}$$

Then,

$$det(A(\phi)) = \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) a_{1,\sigma(1)} \ldots a_{n,\sigma(n)}$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) l_{\phi(1),\sigma(1)} \ldots l_{\phi(n),\sigma(n)}$$

Now, since $\phi \notin \Phi$, there is some $i_0 \neq j_o$ in $\{1, 2, \ldots, n\}$ such that $\phi(i_0) = \phi(j_0)$. Hence, if R_i denote the ith row of $A(\phi)$, then

$$R_{i_0} = (a_{i_0,1}(\phi), \ldots, a_{i_0,n}(\phi))$$

$$= (l_{\phi(i_0),1}, \ldots, l_{\phi(i_0),n})$$

$$(l_{\phi(j_0),1}, \ldots, l_{\phi(j_0),n}) = R_{j_0}$$

Two rows are same implies the determinant, $det(A(\phi)) = 0$. Thus, I_2 can be taken to be zero and I_1 only survives in the sum. Hence,

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \sum_{\phi \in \Phi} \prod_{i=1}^{n} l_{\phi(i),i} l_{\phi(i),\sigma(i)}$$

Now we use the fact that for each $\phi \in \Phi$, it can be uniquely written as $\phi = \lambda \circ \theta$ where $\theta \in \mathcal{S}_n$ and $\lambda \in \Lambda(m,n)$. So,

$$[[L]]^2 = \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \mathcal{S}_n} \prod_{i=1}^{n} l_{\lambda \circ \theta(i),i} l_{\lambda \circ \theta(i),\sigma(i)}$$

$$= \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \mathcal{S}_n} \prod_{i=1}^{n} l_{\lambda(i),\theta^{-1}(i)} l_{\lambda(i),\sigma \circ \theta^{-1}(i)}$$

$$= \sum_{\lambda \in \Lambda(m,n)} \sum_{\theta \in \mathcal{S}_n} \sum_{\sigma \in \mathcal{S}_n} sgn(\sigma) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)} l_{\lambda(i),\sigma \circ \theta(i)}$$

Letting $\rho = \sigma \circ \theta$, we obtain,

$$= \sum_{\lambda \in \Lambda(m,n)} \sum_{\rho \in \mathcal{S}_n} \sum_{\theta \in \mathcal{S}_n} sgn(\theta) sgn(\rho) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)} l_{\lambda(i),\rho(i)}$$

$$= \sum_{\lambda \in \Lambda(m,n)} \left(\sum_{\theta \in \mathcal{S}_n} sgn(\theta) \prod_{i=1}^{n} l_{\lambda(i),\theta(i)} \right)^2$$

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

□

Remark 1.7. To calculate $[[L]]^2$, we compute the sum of squares of the determinants of each $n \times n$ - submatrices of the $m \times n$ - matrix representing L.

Definition 1.8. Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}^m$ be a map. Let $x_0 \in U$. f is said to be fréchet differentiable at x_0 if there exists $A : \mathbb{R}^n \to \mathbb{R}^m$, a linear map such that,

$$\left\{ \begin{array}{c} \text{Given } \epsilon > 0, \exists \delta > 0 \text{ such that} \\ ||f(x_0 + h) - f(x_0) - Ah|| < \epsilon |h| \\ \text{whenever } |h| < \delta \end{array} \right\}$$

Here, $||.||$ denoted the euclidean norm in $\mathbb{R}^m$ and $|.|$ denoted the euclidean modulus in $\mathbb{R}^n$.

Lemma 1.9. *Let $f = (f_1, \dots, f_m)$ be fréchet differentiable, then $\frac{df_i}{dx_j}(x_0)$ exists and*

$$A = \begin{bmatrix} \frac{df_1}{dx_1}(x_0) & \frac{df_1}{dx_2}(x_0) & \dots & \frac{df_1}{dx_n}(x_0) \\ \frac{df_2}{dx_1}(x_0) & \frac{df_2}{dx_2}(x_0) & \dots & \frac{df_2}{dx_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{df_m}{dx_1}(x_0) & \frac{df_m}{dx_2}(x_0) & \dots & \frac{df_m}{dx_n}(x_0) \end{bmatrix}$$

Proof. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

and $e_i = (0, 0, \dots, 1, \dots, 0)$, i.e having 1 at the ith position.
Then

$$Ae_i = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

Let $h = \eta e_i$, then ,

$$Ah = \eta A e_i = \eta \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix}$$

Hence,

$$||f(x_0 + h) - f(x_0) - Ah|| = ||f(x_0 + \eta e_i) - f(x_0) - \eta \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} ||$$

$$= \left\| \begin{matrix} f_1(x_0 + he_i) - f_1(x_0) - \eta a_{1n} \\ f_2(x_0 + he_i) - f_2(x_0) - \eta a_{2n} \\ \vdots \\ f_m(x_0 + he_i) - f_m(x_0) - \eta a_{mn} \end{matrix} \right\| < \epsilon|\eta| \ ; \ \forall|\eta| < \delta$$

$$\implies |f_j(x_0 + \eta e_i) - f_j(x_0) - \eta a_{ji}| < \epsilon|\eta| \ ; \ \forall |\eta| < \delta$$

$$\implies |\frac{f_j(x_0 + \eta e_i) - f_j(x_0)}{\eta} - a_{ji}| < \epsilon \ ; \ \forall |\eta| < \delta$$

$$\implies a_{ji} = \frac{df_j}{dx_i}(x_0)$$

This proves the lemma.

□

Definition 1.10. $f : U \to \mathbb{R}^m$ is said to be fréchet differentiable in U if $\forall x \in U$, $f'(x)$ exists and is given by

$$f'(x) = \begin{bmatrix} \frac{df_1}{dx_1}(x) & \frac{df_1}{dx_2}(x) & \dots & \frac{df_1}{dx_n}(x) \\ \frac{df_2}{dx_1}(x) & \frac{df_2}{dx_2}(x) & \dots & \frac{df_2}{dx_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{df_m}{dx_1}(x) & \frac{df_m}{dx_2}(x) & \dots & \frac{df_m}{dx_n}(x) \end{bmatrix}$$

which is called the gradient matrix of f at x.

Remark 1.11. For f linear map, i.e $f(x) \equiv A.x$, $f'(x)$ is same as A for all x.

Definition 1.12. For $f : \mathbb{R}^n \to \mathbb{R}^m$, assume f to be fréchet differentiable at x, then, for the gradient matrix $Df(x)$, defined as above, which is a $m \times n$ matrix, we define the Jacobian of f at x, denoted by $Jf(x)$, to be

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

Remark 1.13. f is fréchet differentiable at x_0 implies that all the partial derivatives of f exist at x_0.
The converse is not true, in general.
Counter-example : Consider the following function,

$$f(x, y) = \begin{Bmatrix} \frac{xy}{x^2+y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{else} \end{Bmatrix}$$

Here,

$$\frac{df}{dx}(0, 0) = \frac{f(\eta, 0) - f(0, 0)}{\eta} = 0$$

$$\frac{df}{dy}(0, 0) = \frac{f(0, \eta) - f(0, 0)}{\eta} = 0$$

And hence, $(\frac{df}{dx}(0,0), \frac{df}{dy}(0,0))$ exists.
Suppose f is fréchet differentiable at $(0,0)$, then ,

$$f'(0,0) = (\frac{df}{dx}(0,0), \frac{df}{dy}(0,0)) = (0,0)$$

Declare $0^* := (0,0) \in \mathbb{R}^2$, then

$$|f(0^* + h) - f(0^*) - f'(0^*)h| < \epsilon|h| \ ;\forall |h| < \delta$$

$$\implies |f(h)| < \epsilon|h| \ ; \ |h| < \delta$$

Let $h = \eta(1,1) = (\eta, \eta)$. Then $|h| = \sqrt{(2)}|\eta| < \delta \implies |\eta| < \delta/\sqrt{2}$
Thus, $f(h) = f(h_1, h_2) = f(\eta, \eta) = 1$ and this implies $1 = |f(h)| < (\sqrt{2})\epsilon\eta$.
Let $\eta \to 0$ to get a contradiction.

□

Theorem 1.14. *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a map such that*

- $\frac{df_j}{dx_i}(x)$ *exist,* $\forall\ 1 \leq j \leq m, \forall\ 1 \leq i \leq n, \forall\ x \in U$
- $x \to \frac{df_j}{dx_i}(x)$ *is continuous.*

Then f is fréchet differentiable in U and as a $m \times n$ matrix,

$$f'(x) = \left[(\frac{df_j}{dx_i}) \right]_{1 \leq j \leq m, 1 \leq i \leq n}$$

Proof. Denote

$$G(x) = \left[(\frac{df_j}{dx_i}) \right]_{1 \leq j \leq m, 1 \leq i \leq n}$$

By the second point in the hypothesis of the theorem, G is continuous and hence, $\forall x \in U$, given $\epsilon > 0, \exists \delta > 0$ such that

$$||G(x+h) - G(x)|| < \epsilon, \ \forall \ |h| < \delta$$

Now, it is enough to show that $\forall j, f_j : U \to \mathbb{R}$ is differentiable. Hence, consider for the case $m = 1$, i.e $f : U \to \mathbb{R}$ and $(\frac{df}{dx_1}, \dots \frac{df}{dx_n})(x)$ exist and is continuous, as per the hypothesis of the theorem.
Let $x \in U$ and $\delta > 0$ such that $B(x, \delta) \subset U$. Let $|h| < \delta$, then $x + h \in B(x, \delta) \subset U$.

Let $h = (h_1, \dots, h_n)$ and $G(x) = (\frac{df}{dx_1} \cdots \frac{df}{dx_n})(x)$. Consider

$$f(x+h) - f(x) = f(x_1 + h_1, \dots, x_n + h_n) - f(x_1, \dots x_n)$$

$$= \Big[f(x_1 + h_1, x_2 + h_2, x_3 + h_3, \dots, x_n + h_n) - f(x_1, x_2 + h_2, \dots, x_n + h_n)$$

$$+ f(x_1, x_2 + h_2, \dots, x_n + h_n) - f(x_1, x_2, x_3 + h_3 \dots, x_n + h_n)$$

$$+ f(x_1, x_2, x_3 + h_3 \dots, x_n + h_n) - f(x_1, x_2, x_3, \dots, x_n + h_n)$$

$$\vdots$$

$$+ f(x_1, \dots, x_{n-1}, x_n + h_n) - f(x_1, \dots x_n) \Big]$$

$$= \sum_{i=1}^{n} \Big[f(x_1, \dots, x_{i-1}, x_i + h_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n) - f(x_1, \dots, x_{i-1}, x_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n) \Big]$$

Thus,

$$f(x+h) - f(x) - G(x).h$$

$$= \sum_{i=1}^{n} \Big[f(x_1, \dots, x_{i-1}, x_i + h_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n)$$

$$- f(x_1, \dots, x_{i-1}, x_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n)$$

$$- \frac{df}{dx_i}(x_1, \dots, x_n) h_i \Big]$$

$$= \sum_{i=1}^{n} h_i \int_0^1 \Big[\frac{df}{dx_i}(x_1, \dots, x_{i-1}, x_i + t h_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n)$$

$$- \frac{df}{dx_i}(x_1, \dots, x_n) \Big] dt$$

Now, since $\frac{df}{dx_i}$ is continuous, for a given $\epsilon > 0, \exists \delta > 0$ such that

$$\forall |h| < \delta, \ \left| \frac{df}{dx_i}(x_1, \dots, x_{i-1}, x_i + t h_i, x_{i+1} + h_{i+1}, \dots, x_n + h_n) - \frac{df}{dx_i}(x_1, \dots, x_n) \right| < \epsilon$$

$$\implies |f(x+h) - f(x) - G(x).h| \leq \epsilon (\sum_{i=1}^{n} |h_i|), \ \forall \sum |h_i| < \delta$$

This tells that f is fréchet differentiable and $f' = G$

□

Theorem 1.15. *Let $U \subset \mathbb{R}^n$ open, $V \subset \mathbb{R}^m$ open.*
Let $f : U \to V$, $g : V \to \mathbb{R}^p$ be two fréchet differentiable maps. Then $g \circ f : U \to \mathbb{R}^p$ is

fréchet differentiable and as linear maps,

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x)$$

Remark 1.16. For fixed x, $f'(x) : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map and $g'(f(x)) : \mathbb{R}^m \to \mathbb{R}^p$ is a linear map. Hence,

$$g'(f(x)) \circ f'(x) : \mathbb{R}^n \to \mathbb{R}^p$$

is a linear map.
Explicitly, let $f = (f_1, f_2, \dots f_m)$, $g = (g_1, g_2, \dots, g_p)$.
Then, in their respective coordinates,

$$f'(x) = \begin{bmatrix} \frac{df_1}{dx_1}(x) & \frac{df_1}{dx_2}(x) & \dots & \frac{df_1}{dx_n}(x) \\ \frac{df_2}{dx_1}(x) & \frac{df_2}{dx_2}(x) & \dots & \frac{df_2}{dx_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{df_m}{dx_1}(x) & \frac{df_m}{dx_2}(x) & \dots & \frac{df_m}{dx_n}(x) \end{bmatrix}$$

$$g'(y) = \begin{bmatrix} \frac{dg_1}{dy_1}(y) & \frac{dg_1}{dy_2}(y) & \dots & \frac{dg_1}{dy_m}(y) \\ \frac{dg_2}{dy_1}(y) & \frac{dg_2}{dy_2}(y) & \dots & \frac{dg_2}{dy_m}(y) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{dg_p}{dy_1}(y) & \frac{dg_p}{dy_2}(y) & \dots & \frac{dg_p}{dy_m}(y) \end{bmatrix}$$

Hence,

$$g'(f(x)) \circ f'(x) = \begin{bmatrix} \frac{dg_1}{dy_1}(f(x)) & \frac{dg_1}{dy_2}(f(x)) & \dots & \frac{dg_1}{dy_m}(f(x)) \\ \frac{dg_2}{dy_1}(f(x)) & \frac{dg_2}{dy_2}(f(x)) & \dots & \frac{dg_2}{dy_m}(f(x)) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{dg_p}{dy_1}(f(x)) & \frac{dg_p}{dy_2}(f(x)) & \dots & \frac{dg_p}{dy_m}(f(x)) \end{bmatrix} \begin{bmatrix} \frac{df_1}{dx_1}(x) & \frac{df_1}{dx_2}(x) & \dots & \frac{df_1}{dx_n}(x) \\ \frac{df_2}{dx_1}(x) & \frac{df_2}{dx_2}(x) & \dots & \frac{df_2}{dx_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{df_m}{dx_1}(x) & \frac{df_m}{dx_2}(x) & \dots & \frac{df_m}{dx_n}(x) \end{bmatrix}$$

$$= \left[(\sum_{j=1}^{m} \frac{dg_i}{dy_j}(f(x)) \frac{df_j}{dx_l}) \right]_{1 \le i \le p \ ; \ 1 \le l \le n}$$

Proof. Proof of the theorem :
Let

$$f(x+h) - f(x) - f'(x).h = e(h)$$

By the differentiability criteria,

$$\frac{|e(h)|}{|h|} \to 0 \text{ as } |h| \to 0$$

Also, $f(x+h) = f(x) + f'(x)h + e(h)$ holds and hence,

$$g(f(x+h)) = g(f(x) + f'(x)h + e(h))$$

Let

$$\eta = f'(x)h + e(h)$$

Then,

$$g(f(x+h)) = g(f(x) + \eta)$$

So,

$$g(f(x+h)) - g(f(x)) - g'(f(x))f'(x)h$$

$$= g(f(x) + \eta) - g(f(x)) - g'(f(x))f'(x)h$$

$$= g'(f(x))\eta + b(\eta) - g'(f(x))f'(x)h$$

$$= g'(f(x))(\eta - f'(x)h + b(\eta))$$

where $\frac{b(\eta)}{|\eta|} \to 0$ as $|\eta| \to 0$.
Looking at η , $\eta = f'(x)h + e(h) \implies |\eta| \to 0$ as $|h| \to 0$ and

$$|g(f(x+h)) - g(f(x)) - g'(f(x))f'(x)h|$$

$$= |g'(f(x))e(h) + b(\eta)|$$

$$\leq (||g'(f(x))|| + 1)(|e(h)| + |b(\eta)|)$$

$$\text{Thus, } \frac{|g(f(x+h)) - g(f(x)) - g'(f(x))f'(x)h|}{|h|} \leq \frac{||g'(f(x))|| + 1)(|e(h)| + |b(\eta)|)}{|h|}$$

As $|\eta| \to 0$ as $|h| \to 0$, the RHS of the above inequality tends to 0.

Hence $g \circ f$ is differentiable and

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

□

Example 1.1. *let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map given by*

$$T(x) = Tx \equiv \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \dots & t_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then

$$T' = T$$

Proof.

$$T(x+h) - T(x) = T(x) + T(h) - T(x) = T(h)$$

T is linear and by above,

$$T(x+h) - T(x) - T(h) = 0$$

This tells that by the definition of the differential, $T' = T$

$\square$

1.3 Taylor's formula

Notations 1.17.

- *For* $\Omega \subset \mathbb{R}^n$,

$$C^k(\Omega) := \{f : \Omega \to \mathbb{R}\ ;\ f \textit{ is } k \textit{ times differentiable for all points in } \Omega\}$$

- *Closed line segment* $[a, b] := \{ta + (1-t)b; t \in [0, 1]\}$

- *For* $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, *define*

$$D^\alpha := \left(\frac{d}{dx_1}\right)^{(\alpha_1)} \left(\frac{d}{dx_2}\right)^{(\alpha_2)} \cdots \left(\frac{d}{dx_n}\right)^{(\alpha_n)}$$

$$\alpha! := \alpha_1!\alpha_2! \ldots \alpha_n!$$

$$x^\alpha \equiv (x_1, x_2, \ldots, x_n)^\alpha := x_1^{\alpha_1}.x_2^{\alpha_2}.\ldots.x_n^{\alpha_n}$$

$$|\alpha| := \alpha_1 + \alpha_2 \cdots + \alpha_n$$

Theorem 1.18. ***Taylor's formula:***
Let Ω *be open in* $\mathbb{R}^n$ *and* $f \in C^k(\Omega)$. *If* $x, y \in \Omega$ *and the closed line segment* $[x, y]$ *joining* x *to* y *is also in* Ω,*then,*

$$f(x) = \sum_{|\alpha| \le k-1} \frac{D^\alpha f(y)}{\alpha!}(x-y)^\alpha + k \sum_{|\alpha|=k} \frac{(x-y)^\alpha}{\alpha!} \int_0^1 (1-t)^{k-1} D^\alpha f(x + t(y-x))dt$$

Lemma 1.19. *Let N be a neighbourhood of the closed interval $0 \leq t \leq 1$ in $\mathbb{R}$ and let $g \in C^k(N)$, then,*

$$g(1) = \sum_{\beta=0}^{k-1} \frac{g^{(\beta)}(0)}{\beta!} + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} g^{(k)}(t) dt$$

where $g^{(k)}(x)$ denotes the k-th derivative of g at x.

Proof. Integrate by parts, k many times to get :

$$\frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} g^{(k)}(t) dt$$

$$= -\frac{g^{(k-1)}(0)}{(k-1)!} + \frac{1}{(k-2)!} \int_0^1 (1-t)^{k-2} g^{(k-1)}(t) dt$$

$$= -\sum_{j=1}^{k-1} \frac{g^{(j)}(0)}{j!} + \int_0^1 g'(t) dt$$

$$= -\sum_{j=0}^{k-1} \frac{g^{j}(0)}{j!} + g(1) - g(0)$$

Hence,

$$g(1) = \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} g^{(k)}(t) dt$$

Lemma 1.20. *Let $n \geq 1$, $k \geq 0$, $\xi := (\xi_i, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$, with $\xi \in \mathcal{A}^n$, where $\mathcal{A}$ is a commutative algebra, then,*

$$(\xi_1 + \xi_2 + \cdots + \xi_n)^k = k! \sum_{|r|=k} \frac{\xi^r}{r!}$$

Proof. of the lemma : Fix k and the proof is by induction on n. For $n = 2$, by the binomial theorem,

$$(\xi_1 + \xi_2)^k = \sum_{r=1}^{n} \frac{k!}{r!(k-r)!} \xi_1^r \xi_2^{k-r}$$

$$= k! \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!} \text{ ; for } \alpha \equiv (r, k-r)$$

Assume now that the lemma holds upto $n-1$ and let $\eta = \xi_2 + \xi_3 + \cdots + \xi_n$, then,

$$(\xi_1 + \xi_2 + \cdots + \xi_n)^k = (\xi_1 + \eta)^k = k! \sum_{a+b=k} \frac{1}{a!b!} \xi_1^a \eta^b$$

$$= k! \sum_{a+b=k} \frac{\xi_1^a}{a!} \frac{1}{b!} (\xi_2 + \dots \xi_n)^b$$

$$= k! \sum_{a+b=k} \frac{\xi_1^a}{a!} \frac{1}{b!} b! \sum_{\beta_2+\dots+\beta_n=b} \frac{\xi_2^{\beta_1} \dots \xi_n^{\beta_n}}{\beta_2! \dots \beta_n!}$$

$$= k! \sum_{a+b=k} \sum_{\beta_2+\dots+\beta_n=b} \frac{\xi_1^a \xi_2^{\beta_2} \dots \xi_n^{\beta_n}}{a! \beta_1! \dots \beta_n!}$$

$$= k! \sum_{|\gamma|=k} \frac{\xi^n}{\gamma!}$$

This proves the lemma.

□

Corollary 1.21. *Let*

$$L = (\xi_1 \frac{\partial}{\partial x_1} + \dots + \xi_n \frac{\partial}{\partial x_n})$$

be the first order differential operator . Then

$$L^k = k! \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!} D^\alpha$$

Proof. Let

$$\eta = (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$$

Apply the lemma to $(\xi_1 \eta_1 + \dots + \xi_n \eta_n)$ to get :

$$L^k = (\xi_1 \eta_1 + \dots + \xi_n \eta_n)^k$$

$$= k! \sum_{|\alpha|=k} \frac{\xi_1^{\alpha_1} \dots \xi_n^{\alpha_n} \eta_1^{\alpha_1} \dots \eta_n^{\alpha_n}}{\alpha_1! \dots \alpha_n!}$$

$$= k! \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!} D^\alpha$$

□

Proof. of the theorem. Let $x, y \in \Omega$ such that $ty + (1-t)x \in \Omega$ for all $t \in [0,1]$. Let

$$g(t) = f(ty + (1-t)x)$$

Then, $g(0) = f(x), g(1) = f(y)$. Let $\xi_i = y_i - x_i$, then,

$$\frac{dg}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} (x + t(y-x))(y_i - x_i)$$

$$= (\xi_1 \frac{\partial}{\partial x_1} + \dots \xi_n \frac{\partial}{\partial x_n}) f(x + t(y - x))$$

Hence, from the corollary, we have,

$$\frac{d^{(j)} g}{dt^j} = (\xi_1 \frac{\partial}{\partial x_1} + \dots \xi_n \frac{\partial}{\partial x_n})^j f(a + t(y - x))$$

$$= j! \sum_{|\alpha| = j} \frac{\xi^\alpha}{\alpha!} D^\alpha f(a + t(y - x))$$

From the lemma 1.19 , we have,

$$f(y) = \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} + \frac{1}{(k-1)!} \int_0^1 g^{(k)}(t)(1-t)^{k-1} dt$$

$$= \sum_{|\alpha| \leq k-1} \frac{D^\alpha f(x)}{\alpha!} (y - x)^\alpha + \frac{k!}{(k-1)!} \int_0^1 \frac{1}{k!} g^{(k)}(t)(1-t)^{k-1} dt$$

$$= \sum_{|\alpha| \leq k-1} \frac{D^\alpha f(x)}{\alpha!} (y - x)^\alpha + k \sum_{|\alpha| = k} \frac{(y - x)^\alpha}{\alpha!} \int_0^1 (1-t)^{k-1} D^\alpha f(x + t(y - x)) dt$$

This proves the taylor's formula.

□

1.4 Some pre-requisite calculus.

1.4.1 Semi-continuity

Definition 1.22. Let $(X, ||.||)$ be a normed space and $f : X \to \mathbb{R}$ be a function. f is called lower semi-continuous at y if

$$\forall \epsilon, \exists \delta > 0 \; ; ||x - y|| < \delta \implies f(x) > f(y) - \epsilon$$

Definition 1.23. Let $(X(, ||.||)$ be a normed space and $f : X \to \mathbb{R}$ be a function. f is called upper semi-continuous at y if

$$\forall \epsilon, \exists \delta > 0 \; ; ||x - y|| < \delta \implies f(x) < f(y) + \epsilon$$

Theorem 1.24. *f is lower semi-continuous at y $\iff$ for all $x_n \to y$, $\liminf_{n \to \infty} f(x_n) \geq f(y)$.*

Proof. Let f be lower semi- continuous and let $y \in X$. Let $x_n \to y$, that is, $||x_n - y|| \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be chosen and fixed for the argument.
Then, there is $N \in \mathbb{N}$ such that, for $n \geq N$, $||x_n - y|| < \delta$, where δ is chosen as per the definition of lower semi-continuity, i.e $\exists \delta$, for the fixed ϵ, with $f(x_n) > f(y) - \epsilon$. Thus for all $n > N$, $f(x_n) > f(y) - \epsilon$ and hence,

$$\inf_{k \geq N} f(x_k) > f(y) - \epsilon$$

$$\implies \lim_{n \to \infty} \inf_{k \geq n} f(x_k) \geq f(y) - \epsilon$$

$$\implies \liminf_{n \to \infty} f(x_n) \geq f(y)$$

Coversely, suppose that f is not lower semi-continuous at y, then, there is $\epsilon > 0$ such that for any $\delta > 0$, there is x_δ with

$$||x_\delta - y|| < \delta \implies f(x_\delta) \leq f(y) - \epsilon$$

Let $\delta = \frac{1}{n}$ and denote $x_n = x_\delta$.

$$\implies \liminf_{n \to \infty} f(x_n) \leq f(y) - \epsilon \leq f(y)$$

There is $N \in \mathbb{N}$, such that for all $n \geq N$, $||x_n - y|| < \delta$ and thus,

$$\implies \inf_{n > N} f(x_n) \leq f(y) - \epsilon$$

$$\implies \liminf_{n} f(x_n) \leq f(y) - \epsilon$$

This contradicts the fact that :

$$\liminf_{n \to \infty} f(x_n) \geq f(y)$$

□

Remark 1.25. f is upper semi-continuous at $y \iff$ for all $x_n \to y$, $\liminf_{n \to \infty} f(x_n) \leq f(y)$.
The proof is similar to that of the previous theorem, but it has to be applied to the class upper semicontinuous functions.

Theorem 1.26. *f is lower semi- continuous $\iff \forall t$, $U_t := \{x \in X \ ; \ f(x) > t\}$ is open.*

Proof. Let f be lower semi-continuous and let $y \in U_t$, then $f(y) > t$. For ϵ suffeciently small, we have $f(y) > t + \epsilon$. By the assumption , there is $\delta > 0$ such that

$$||x - y|| < \delta \implies f(x) > f(y) - \epsilon > t$$

Thus,

$$B(y, \delta) \subset U_t$$

Thus U_t is open.
Conversely, suppose that f is not lower semi-continuous at y, then, there is an $\epsilon > 0$ so that $\forall \delta > 0$, there is x_δ such that $||x_\delta - y|| < \delta$ and $f(x_\delta) \leq f(y) - \epsilon$. Denote $\delta = \frac{1}{n}$, $x_\delta = x_n$ and $t = f(y) - \frac{\epsilon}{2}$. Then there is a sequence $\{x_n\}$ such that $x_n \to y$ with $f(x_n) \leq f(y) - \epsilon = 1 - \frac{\epsilon}{2}$. But $f(y) = t + \frac{\epsilon}{2} > t \implies y \in U_t$. Thus, there is some $r > 0$ such that $B(y, r) \subset U_t$. Now, $x_n \to y$ as $n \to \infty$ and hence $\exists N \in \mathbb{N}$ such that, for all $n > N$, $||x_n - y|| < r$ and $f(x_n) > t$. This contradicts the fact that $f(x_n) \leq t - \frac{\epsilon}{2}$.

□

Remark 1.27. f is upper semi- continuous $\iff \forall t, U_t := \{x \in X \ ; \ f(x) < t\}$ is open. The proof is similar to the above theorem.

Remark 1.28. This theorem tells that f is either upper or lower semi-continuous implies that it is measurable, where the definition of measurable functions is given in the next section.

1.5 Measure Theory

1.5.1 Some set theoretic measure theory.

Let X be a non empty set and $\mathbb{P}(X)$ denote the set of all subsets of X.

Definition 1.29. $m : \mathbb{P}(X) \to \mathbb{R}^+ \equiv [0, \infty)$ is called a measure if for all $A, B, A_i \in \mathbb{P}(X)$ $(\ ; \ i = 1, 2, \dots)$, m satisfies

- $m(\phi) = 0$
- Monotone property : If $A \subset B$, then $m(A) \leq m(B)$
- Countable sub-additivity : If $A = \bigcup_{i=1}^{\infty} A_i$, then $m(A) \leq \sum_{i=1}^{\infty} m(A_i)$

Here (X, m) is called a measure space.

Definition 1.30. Consider (X, m) to be the measure space as defined in the earlier point. A set $A \subset X$ is called measurable set (or m - measurable set) if for all $E \subset X$,

$$m(E) = m(E \cap A) + m(E \cap A^c)$$

Definition 1.31. A measure m on X is called regular if for all sets $A \subset X$, there is a $m-$ measurable set B such that $A \subset B$ and $m(A) = m(B)$

Definition 1.32. A measure m on $\mathbb{R}^n$ is called borel if every borel set is $m-$ measurable,i.e every borel set B satisfies

$$\forall A \subset \mathbb{R}^n \ , \ m(A) = m(A \cap B) + m(A \cap B^c)$$

Definition 1.33. A measure m is borel regular if

- m is Borel
- For all $A \subset \mathbb{R}^n$ there is a borel set B such that $A \subset B$ and $m(A) = m(B)$

Example : For a measure on a topological space to be borel, but not borel regular. Consider the space to be $\mathbb{R}$ with the topology given by just $\{\phi, \mathbb{R}\}$. Associate counting measure m to the above space. Observe that since, $m(\mathbb{R}) = \infty$, $m(\phi) = 0$, $m(\{0\}) = 1$, there is no borel set B, open in the prescribed topology such that $m\{0\} \equiv mB$, in measure.

Example : For a measure on a topological space such that , there is some borel set such that it cannot be approximated by open sets.
Consider the space to be $\mathbb{R}$ and associate it with the counting measure m along with the standard topology, generated by open intervals. Note that any open sets can be written as disjoint union of open intervals. Since, each intervals contain infinitley many point, counting measure of any open set is by default ∞. Also, clearly $\{0\}$ is a borel set. As

$$m(\{0\}) = \inf\{m(U) \ ; \ \{0\} \subset U \ , \ U \text{ is open.}\}$$

LHS = 1. But RHS is always infinity. Hence, $\{0\}$, which is a borel set, cannot be approximated by open sets from the outside in measure.

Definition 1.34. A measure m on $\mathbb{R}^n$ is a radon measure if m is borel regular and $m(K) < \infty$ for all compact sets $K \subset \mathbb{R}^n$

Definition 1.35. For X a set and $\mathbb{P}(X)$ to be it's power set, we define $\mathcal{F} \subset \mathbb{P}(X)$ to be sigma-algebra (σ - algebra) of X if

- $\phi, X \in S$
- $\{A_i\}_{i=1}^{\infty} \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$
- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

Definition 1.36. For X a set and $\mathbb{P}(X)$ to be it's power set, we define $\mathcal{S} \subset \mathbb{P}(X)$ to be semi-algebra of X if

- $\phi, X \in S$
- $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$
- $A \in \mathcal{S} \implies A^c = \bigcup_{finite} B_i$; such that B_i's are disjoint and are elements of $\mathcal{S}$

Definition 1.37. For the before defined X and $\mathbb{P}(X)$, $\mathcal{A} \subset \mathbb{P}(X)$ is called an algebra of sets if

- $\mathcal{A}$ is a semi algebra.
- $A, B \in \mathcal{A} \implies A \bigcup B \in \mathcal{A}$

Theorem 1.38. ***Caratheodory theorem:***
Let (X, m) be a measure space and define $A \in \mathbb{P}(X)$ is said to be measurable if for all $E \in \mathbb{P}(X)$,

$$m(E) = m(E \cap A) + m(E \cap A^c)$$

Let

$$\mathcal{F} = \{ \text{ all measurable sets as defined earlier} \}$$

Then

- *$\mathcal{F} \neq \phi$.*
- *$\mathcal{F}$ is a $\sigma-$ algebra.*
- *Countable additivity : If $A = \bigcup_{i=1}^{\infty} A_i$, with $A, A_i \in \mathcal{F}$, A_i's are disjoint, then*

$$m(A) = \sum_{i=1}^{\infty} m(A_i)$$

- *Complete measure space : For a set $N \in \mathbb{P}(X)$, it is called a null set if $m(N) = 0$. Here all the null sets belong to $\mathcal{F}$ for the above definiton of $m, \mathcal{F}$.*

Lemma 1.39. *If A and B are in $\mathcal{F}$, then $A \bigcup B, A^c, B^c \in \mathcal{F}$, i.e if A and B are measurable, then $A \bigcup B$ and A^c, B^c is measurable.*

Proof. Clearly, $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$, by the definition of measurable sets. Any set (in particular , here A) is measurable implies for any $C \in \mathbb{P}(X)$, we have

$$m(C) = m(C \cap A) + m(C \cap A^c)$$

Notice that it is sufficient that $m(C) \geq m(C \cap A) + m(C \cap A^c)$ holds, is equivalent to saying that the set A is measurable, as the other inequality follows from sub additive property. In particular, we have for $C \equiv C \cap B^c$,

$$m(C \cap B^c) = m(C \cap B^c \cap A) + m(C \cap B^c \cap A^c)$$

Notice that

$$C \cap (A \cup B) = (C \cap B) \cup (C \cap A \cap B^c)$$

Hence,

$$m(C \cap (A \cup B)) \leq m(C \cap B) + m(C \cap A \cap B^c)$$

Thus, from the above 2 equations,

$$m(C \cap (A \cup B)) + m(C \cap B^c \cap A^c) \leq m(C \cap B) + m(C \cap A \cap B^c) + m(C \cap A^c \cap B^c)$$

$$= m(C \cap B) + m(C \cap B^c)$$

$$= m(C)$$

This proves the lemma. $\square$

Lemma 1.40. *For any set $A \in \mathbb{P}(X)$ and $E_1, E_2, \ldots, E_n$, finite sequence of disjoint measurable sets , we have*

$$m\Big(A \cap [\bigcup_{i=1}^{n} E_i]\Big) = \sum_{i=1}^{n} m(A \cap E_i)$$

Proof. The proof proceeds by the induction on n. This is clearly true for $n = 1$. Assume that the statement is true upto the case $n-1$. For the case of n, notice that E_i's are disjoint and hence,

$$A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n = A \cap E_n$$

$$A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n^c = A \cap \left[\bigcup_{i=1}^{n-1} E_i\right]$$

Hence,

$$m\Big(A \cap [\bigcup_{i=1}^{n} E_i]\Big) = m(A \cap E_n) + m\Big(A \cap [\bigcup_{i=1}^{n-1} E_i]\Big)$$

By the process of induction, we have

$$= m(A \cap E_n) + \sum_{i=1}^{n-1} m(A \cap E_i)$$

$$= \sum_{i=1}^{n} m(A \cap E_i)$$

This proves the lemma. □

Proof. of the theorem 1.38:

- $\mathcal{F}$ is not empty as trivially $X, \phi \in \mathcal{F}$ i.e they satisfy the formula for the measurability of sets, as mentioned in the hypothesis of the theorem.

- $\mathcal{F}$ is a sigma algebra :

  - Clearly, $X, \phi \in \mathcal{F}$.
  - Lemma 1.39 tells that $\mathcal{F}$ is an algebra of sets. If $\{A_i\}_{i=1}^{\infty} \in \mathcal{F}$, then it is required to show $\bigcup_{i=1}^{\infty} A_i$ is in $\mathcal{F}$.
    Declare
    $$B_n = \left\{ \begin{array}{cc} A_n & \text{if n = 1.} \\ A_n - \left[\bigcup_{i=1}^{n-1} A_i\right] & else. \end{array} \right\}$$
    Hence,
    $$\bigcup_i A_i = \bigcup_i B_i \text{ , say , it is equal to } E$$
    and
    B_i's are pairwise disjoint, by construction.

    Let
    $$G_n = \bigcup_{i=1}^{n} B_i$$
    Then G_n's are measurable for all n and $E^c \subset G_n^c$ for all n.
    So, for any arbitrary set A, we have

    $$m(A) = m(A \cap G_n) + m(A \cap G_n^c) \geq m(A \cap G_n) + m(A \cap E^c)$$

By the lemma 1.40,

$$m(A \cap G_n) = \sum_{i=1}^{n} m(A \cap B_i)$$

Thus, combining the above 2 equations , we have

$$m(A) \geq \sum_{i=1}^{n} m(A \cap B_i) + m(A \cap E^c)$$

Note that the LHS of the above is independent of n and hence,

$$m(A) \geq \sum_{i=1}^{\infty} m(A \cap B_i) + m(A \cap E^c)$$

$$\geq m(A \cap E) + m(A \cap E^c)$$

This proves that the countable union of measurable sets is measurable.

– If A is measurable, i.e it satisfies the formula mentioned in the hypothesis, A^c also satisfy the same formula as for A and hence $A^c \in \mathcal{F}$.

- To prove the additivity of measure of countable disjoint sets $\{A_i\}$.

$$\bigcup_{i=1}^{n} A_i \subset \bigcup_{i=1}^{\infty} A_i$$

Hence,

$$m\Big(\bigcup_{i=1}^{\infty} A_i\Big) \geq m\Big(\bigcup_{i=1}^{n} A_i\Big)$$

With the use of $A = X$ in lemma 1.40,

$$= \sum_{i=1}^{n} m(A_i)$$

The left hand side is independent of n. And hence,

$$m(\bigcup_{i=1}^{\infty} A_i) \geq \sum_{i=1}^{\infty} m(A_i)$$

The other inequality follows form the subadditivity.

- If N is a set of measure 0, then for any $A \in \mathbb{P}(X)$, $m(N \cap A) \leq m(N) = 0 \implies m(N \cap A) = 0$. Hence,

$$m(N \cap A) + m(A \cap N^c) = m(A \cap N^c) \leq m(A)$$

Hence all null sets are measurable.

□

Lemma 1.41. *Let $\mathcal{S} \subset \mathbb{P}(X)$ be a semi algebra. Let*

$$\mathcal{A} = \left\{ \bigcup_{i=1}^{k} A_i ; A_i \in \mathcal{S} \right\}$$

Then $\mathcal{A}$ is an algebra.

Proof. $A, B \in \mathcal{A} \implies A = \bigcup_{i=1}^{k} A_i$ and $B = \bigcup_{j=1}^{l} B_j$. Hence,

- $A \cap B = \bigcup_{i,j} (A_i \cap B_j) \in \mathcal{A}$ as $\mathcal{S}$ is closed under finite intersection.
- $A \in \mathcal{S} \implies A^c = \bigcup_{i=1}^{k} A_i$, $A_i \in \mathcal{S}$, by the property of semi algebra and hence $A^c \in \mathcal{A}$.
- Let $A = \bigcup_{i=1}^{k} A_i \in \mathcal{A}$, then $A^c = \cap_{i=1}^{k} A_i^c ; A_i^c \in \mathcal{A} \implies A^c \in \mathcal{A}$

Hence $\mathcal{A}$ is an algebra.

□

Definition 1.42. $\mathcal{A}$ is called the algebra generated by $\mathcal{S}$, constructed as in the previous lemma.

Theorem 1.43. *Let m be a regular measure on X. If*

$$A_1 \subset A_2 \subset \cdots \subset A_n \subset \ldots$$

then

$$\lim_{k \to \infty} m(A_k) = m(\bigcup_{k=1}^{\infty} A_k)$$

Proof. Since m is regular, there are measurable sets $\{C_k\}_{k=1}^{\infty}$ with $A_k \subset C_k$ and $m(A_k) = m(C_k)$ for all k. Declare

$$B_k = \cap_{j \geq k} C_j$$

Then, $A_k \subset B_k \subset C_k$ and each B_k is $m-$ measurable. Also, $m(A_k) = m(B_k)$. Thus,

$$\lim_{k \to \infty} m(A_k) = \lim_{k \to \infty} m(B_k) = m\left(\bigcup_{k=1}^{\infty} B_k \right) \geq m\left(\bigcup_{k=1}^{\infty} A_k \right)$$

Note that $A_k \subset \bigcup_{j=1}^{\infty} A_j$. Hence,

$$\lim_{k\to\infty} m(A_k) \leq m\left(\bigcup_{k=1}^{\infty} A_k\right)$$

□

Remark 1.44. Here A_k's need not be measurable.

Definition 1.45. Let m be a measure on X and $A \subset X$. Then m restricted to A, written as

$$(m|A)$$

is the measure defined by

$$(m|A)(B) := m(A \cap B) \text{ ; for all } B \subset X$$

Lemma 1.46. *Let m be a borel, regular measure on $\mathbb{R}^n$. Suppose $A \subset \mathbb{R}^n$ is $m-$ measurable and $m(A) < \infty$. Then $(m|A)$ is radon measure.*

Proof. Declare

$$\overline{m} = (m|A)$$

Since every $m-$ measurable set is $\overline{m}-$ measurable, $\overline{m}$ is a borel measure.
Claim: $\overline{m}$ is borel regular.
Proof of the claim : Since m is borel regular, there exist a borel set B such that $A \subset B$ and $m(A) = m(B) < \infty$. Since A is $m-$ measurable,

$$0 = m(B) - m(A) = m(B - A)$$

Choose $C \subset \mathbb{R}^n$, then,

$$(m|B)(C) = m(C \cap B)$$

$$= m(C \cap B \cap A) + m((C \cap B) - A)$$

$$\leq m(C \cap A) + m(B - A)$$

$$= (m|A)(C)$$

Thus, by subadditivity and above result,

$$(m|B) = (m|A)$$

And hence, one can assume A is a borel set.
Consider $C \subset \mathbb{R}^n$. It is required to show the existence of a borel set D such that

- $C \subset D$
- $\overline{m}(C) = \overline{m}(D)$

Since, m is borel, regular measure, there is a borel set E such that

$$A \cap C \subset E$$

$$m(E) = m(A \cap C)$$

Declare

$$D := E \cup (\mathbb{R}^n - A)$$

Since A and E are borel, so is D. Also,

$$C \subset (A \cap C) \cup (\mathbb{R}^n - A) \subset D$$

Finally, since $D \cap A = E \cap A$ holds, we now have

$$\overline{m}(D) = m(D \cap A) = m(E \cap A) \leq m(E) = m(A \cap C) = \overline{m}(C)$$

Clearly $\overline{m}(K) < \infty$, for all compact sets K.

□

Remark 1.47. If A is a borel set, then $(m|A)$ is borel regular, irrespective of finiteness of $m(A)$.

Lemma 1.48. *Let (X, d) be a metric space and $\overline{m}$ be a measure on X.*
Let $\mathcal{H}$ denote the set of all $\overline{m}-$ measurable sets.
(A) Assume that

- $\overline{m}(X) < \infty$
- $\mathcal{B} \subset \mathcal{H}$, *where $\mathcal{B}$ is the $\sigma-$ algebra of borel sets.*

Let $A \in \mathcal{B}$ and $\epsilon > 0$, then there is a closed set F and an open set U such that

- $F \subset A \subset U$
- $\overline{m}(A - F) \leq \frac{\epsilon}{2}$
- $\overline{m}(U - A) < \frac{\epsilon}{2}$

(B) Assume that

- $\overline{m}$ *is borel.*
- *There is a sequence of open sets* $\{U_k\}_{k=1}^{\infty}$ *such that*
  - $\overline{m}(U_k) < \infty$
  - $X = \bigcup_{k=1}^{\infty} U_k$

  Let A *be a borel set, then*

$$\overline{m}(A) \equiv \inf\{\overline{m}(U) \ ; \ A \subset U, U \ is \ open\}$$

Proof. (A) : Let

$$\mathcal{F} = \Big\{A \in \mathcal{H} \ ; \text{ For each } \epsilon > 0, \text{ there is a closed set } C \subset A \text{ such that } \overline{m}(A - C) < \epsilon\Big\}$$

Notice that, by definition, all closed sets are in $\mathcal{F}$.
Claim 1: If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, then $A := \cap_{i=1}^{\infty} A_i \in \mathcal{F}$.
Proof of the claim 1 : Fix $\epsilon > 0$. Since, $A_i \in \mathcal{F}$, there is a closed set $C_i \subset A_i$ with

$$\overline{m}(A_i - C_i) < \frac{\epsilon}{2^i} \ (i = 1, 2, \dots)$$

Let $C := \cap_{i=1}^{\infty} C_i$. Then C is clearly closed and

$$\overline{m}(A - C) = \overline{m}(\cap_{i=1}^{\infty} A_i - \cap_{i=1}^{\infty} C_i)$$

$$\leq \overline{m}(\bigcup_{i=1}^{\infty}(A_i - C_i))$$

$$\leq \sum_{i=1}^{\infty} \overline{m}(A_i - C_i) < \epsilon$$

And hence, $A \in \mathcal{F}$. This proves the claim 1.
Claim 2 : If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, then $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.
Proof of the claim 2 : Fix $\epsilon > 0$. Choose C_i as above and since $\overline{m}(A) \leq \overline{m}(X) < \infty$,

$$\lim_{m \to \infty} \overline{m}(A - \bigcup_{i=1}^{m} C_i) = \overline{m}(\bigcup_{i=1}^{\infty} A_i - \bigcup_{i=1}^{\infty} C_i)$$

$$\leq \overline{m}(\bigcup_{i=1}^{\infty}(A_i - C_i))$$

$$\leq \sum_{i=1}^{\infty} \overline{m}(A_i - C_i) < \epsilon$$

Thus, by the convergence of the tail of the series to 0, there is an integer m such that

$$\overline{m}(A - \bigcup_{i=1}^{m} C_i) < \epsilon$$

But $\bigcup_{i=1}^{m} C_i$ is closed and hence, $A \in \mathcal{F}$. This proves the claim 2.
Claim 3 : Every open set of X, can be written as the countable union of closed sets.
Proof of the claim 3 : If $\partial U = \phi$, then $U = X$ which is both open and closed. Hence, consider the case $\partial U \neq \phi$. Let

$$F_n : \{x \in U \ ; \ d(x, \partial U) \geq \frac{1}{n}\}$$

F_n's are clearly closed and $U = \bigcup_{n=1}^{\infty} F_n$.
This proves the claim 3.
Thus, by the claim 1, $U \in \mathcal{F}$. By the claim 3, we have $\mathcal{F}$ contains all the open sets as well. Now declare

$$\mathcal{G} := \{A \in \mathcal{F} \ ; \ \mathbb{R}^n - A \in \mathcal{F}\}$$

The purpose of construction of $\mathcal{G}$ is that if $A \in \mathcal{G}$, then $A^c \in \mathcal{G}$. Also, notice that $\mathcal{G}$ contains all the open sets.
Claim 4 : If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{G}$, then $A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{G}$.
Proof of the claim 4 : Claim 4 is a trivial consequence of the claim 2 and the fact of demovier's inequality that is $(\bigcup A_i)^c = \cap(A_i^c)$ and $(\cap A_i)^c = \bigcup(A_i^c)$. This proves the claim 4.
Thus $\mathcal{G}$ is a sigma algebra containing all the open sets and by the definition of the borel sigma algebra, i.e the smallest sigma algebra containing the open sets, $\mathcal{G}$ contains the borel sigma algebra. In particular, $B \in \mathcal{G}$ and hence by the construction of $\mathcal{F}$ and $\mathcal{G}$, given $\epsilon > 0$, there is a closed set $C \subset B$ such that

$$\overline{m}(B - C) < \epsilon$$

Now, $C \subset X - E \iff E \subset X - C = U$, where U is open.

$$\implies \overline{m}(U - E) = \overline{m}\big((X - C) - E\big) = \overline{m}\big((X - E) - C\big) < \epsilon$$

$$\implies \overline{m}(E) = \inf\{m(U) \ ; \ U \text{ is open and } E \subset U\}$$

This proves (A).
(B) : Let D be a borel set and $m_n := (\overline{m}|U_n)$, then m_n is a finite borel measure and hence, by the earlier part of this lemma, for every $\epsilon > 0$, there is an open set V_n' such that

- $U_n \cap D \subset V_n'$
- $\overline{m}\big((U_n \cap V_n') - (U_n \cap D)\big) < \frac{\epsilon}{2^n}$

Declare $V_n := V_n' \cap U_n$, then

$$U_n \cap D \subset V_n \subset V_n'$$

$$V_n \subset U_n$$

$$\overline{m}\big(V_n - (U_n \cap D)\big) = \overline{m}\big((V_n' \cap U_n) - (U_n \cap D)\big) < \frac{\epsilon}{2^n}$$

Let $V = \bigcup_{n=1}^{\infty} V_n$, then

$$D = \bigcup_{n=1}^{\infty} U_n \cap D \subset V$$

$$\overline{m}(V - D) = \overline{m}\left(\bigcup_{n=1}^{\infty} V_n - \bigcup_{n=1}^{\infty}(U_n \cap D)\right) \leq \sum_{i=1}^{\infty} \overline{m}\Big(V_i - (U_i \cap D)\Big) < \epsilon$$

$$\implies \overline{m}(D) = \inf\{\overline{m}(V) \ ; \ D \subset V\}$$

This proves the (B) part and hence the lemma. □

Corollary 1.49. *Let m be a borel measure on $\mathbb{R}^n$ and A be a borel set such that $m(A) < \infty$. Then,*

- $$m(A) = \sup\{m(F) \ ; \ F \subset A, F \textit{ is closed. } \}$$

  $$m(A) = \inf\{m(U) \ ; \ A \subset U, U \textit{ is open. } \}$$

- *Let m be a radon measure. Then, for all sets $A \subset \mathbb{R}^n$,*

  $$m(A) = \inf\{A \subset U \ ; \ U \textit{ is open. } \}$$

Proof. Identify $\mathbb{R}^n$ as a metric space with the usual euclidean metric, i.e

$$d(x, y) \equiv d\big((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)\big) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

Let $m_1(C) = m(A \cap C)$, then, from the previous lemma, (1) follows. Let m be a radon measure. Then, for every compact set $K \subset \mathbb{R}^n$, $m(K) < \infty$.

$$\implies m(U_k) = m\big(B(0, k)\big) \leq m\big(\overline{B(0, k)}\big) < \infty$$

where $U_k := B(0, k)$ is a ball around 0 with radius k. Note that $\mathbb{R}^n = \bigcup_{k=1}^{\infty} B(0, k)$. Let $A \subset \mathbb{R}^n$. Case 1 : $m(A) = \infty$: Take $U = \mathbb{R}^n$.

Case 2 : $m(A) < \infty$. By the borel regularity, there is a borel set E such that $A \subset E$ and $m(A) = m(E) < \infty$. From the previous lemma,

$$m(A) = m(E) = \inf\{m(U) \ ; \ E \subset U, U \text{ is open. }\}$$

Since $A \subset W$, W is open, it implies that $m(A) \leq m(W)$ and thus,

$$m(A) \leq \inf\{m(W) \, ; \, W \text{ is open}, A \subset W\} \leq \inf\{m(U) \, ; \, E \subset U, U \text{ is open }\} = m(E) = m(A)$$

This proves the corollary.

□

Theorem 1.50. ***Caratheodory's criterion:***
Let m be a measure on $\mathbb{R}^n$ as defined earlier. If $m(A \bigcup B) = m(A) + m(B)$ for all $A, B \subset \mathbb{R}^n$ with $dist(A, B) := inf\{|a - b| \ ; \ a \in A, b \in B\}$ to be strictly positive, then m is borel measure.

Proof. It is sufficient to show that all closed sets are measurable. Suppose $C \subset \mathbb{R}^n$ is closed, it is required to show that for any $A \subset \mathbb{R}^n$,

$$m(A) \geq m(A \cap C) + m(A - C)$$

If $m(A) = \infty$, then the above inequality is obvious. Now assume $m(A) < \infty$.
Declare

$$C_k = \{x \in \mathbb{R}^n \ ; \ dist(x, C) \leq \frac{1}{k}\} \ ; \text{for all } k \in \mathbb{N}$$

Then, $dist(A - C_k, A \cap C) \geq \frac{1}{k} > 0$. From the hypothesis, we have

$$m(A - C_k) + m(A \cap C) = m((A - C_k) \cup (A \cap C)) \leq m(A)$$

Claim : $\lim_{n \to \infty} m(A - C_n) = m(A - C)$.
Suppose that the claim is true, then taking limit $n \to \infty$ on both the sides, we get,

$$m(A - C) + m(A \cap C) \leq m(A)$$

This proves the theorem.
Proof of the claim : Declare

$$R_k := \{x \in A \ ; \ \frac{1}{k+1} < dist(x, C) \leq \frac{1}{k}\} \ ; \text{for all } k \in \mathbb{N}$$

Let $x \in A - C$. Suppose that $d(x, C) \leq \frac{1}{k}$, for all k, then $d(x, C) = 0$. Hence, $x \in \overline{C} = C$, which is a contradiction.

Thus $d(x, C) > 0$.
Now, for $k \neq l$,

$$R_{2k} \cap R_{2l} = \phi$$

$$R_{2k+1} \cap R_{2l+1} = \phi$$

$$\implies \sum_{l=1}^{k} m\big(R_{2l}\big) = m\Bigg(\bigcup_{l=1}^{k} R_{2l}\Bigg) \leq m(A)$$

and

$$\sum_{l=1}^{k} m\big(R_{2l+1}\big) = m\Bigg(\bigcup_{l=1}^{k} R_{2l+1}\Bigg) \leq m(A)$$

$$\implies \sum_{l=1}^{2k+1} m(R_l) \leq \sum_{l=1}^{k} m(R_{2l}) + \sum_{l=1}^{k} m(R_{2l+1}) \leq 2m(A) < \infty$$

Letting $k \to \infty$, we get

$$\sum_{l=1}^{\infty} m(R_l) \leq m(A) < \infty$$

$$\implies \lim_{k\to\infty} \sum_{l=k}^{\infty} m(R_l) = 0$$

Now, for $k > 0$,

$$A - C_k \subset A - C \subset \big(A - C_k\big) \cup \bigcup_{l=k}^{\infty} R_l$$

$$\implies m\big(A - C_k\big) \leq m(A - C) \leq m\big(A - C_k\big) + \sum_{l=k}^{\infty} m(R_l)$$

Letting $k \to \infty$,

$$\implies \limsup_{k\to\infty} m\big(A-C_k\big) \leq m(A-C) \leq \liminf_{k\to\infty} m\big(A-C_k\big) + \lim_{k\to\infty} \sum_{l=k}^{\infty} m(R_l) = \liminf_{k\to\infty} m(A-C_k)$$

This proves the claim.

□

Definition 1.51. Let $S \subset \mathbb{P}(X)$ be a semi algebra and $m : \mathcal{S} \to \mathbb{R}^+$ is called a measure if

- $m(A) \geq 0$, for all $A \in \mathcal{S}$.
- $A \subset B \implies m(A) \leq m(B)$.
- If $A = \bigcup_{i=1}^{\infty} A_i$ with $A, A_i \in \mathcal{S}$, A_i's are disjoint , then $m(A) = \sum_{i=1}^{\infty} m(A_i)$.

Lemma 1.52. *Let m be a measure on the semi algebra $\mathcal{S} \subset \mathbb{P}(X)$. Let $\mathcal{A}$ be the algebra generated by $\mathcal{S}$. Define,*

$$\overline{m} : \mathcal{A} \to \mathbb{R}^+$$

$$\overline{m}(A) = \sum_{i=1}^{k} m(A_i); \text{ for } A = \bigcup_{i=1}^{k} A_i, \text{ where } A_i\text{'s are disjoint elements of } \mathcal{A}.$$

Then $\overline{m}$ is well defined and satisfies

- $\overline{m}(A) \geq 0$ *for all* $A \in \mathcal{A}$.
- $A = \bigcup_{i=1}^{\infty} A_i$ *with* $A, A_i \in \mathcal{A}$, A_i*'s are disjoint. Then* $\overline{m}(A) = \sum_{i=1}^{\infty} \overline{m}(A_i)$.

Proof. (Step 1) : Let $A = \bigcup_{i=1}^{n} B_i$; $B_i \in \mathcal{S}$. Define

$$C_1 = B_1$$

$$C_2 = B_1^c \cap B_2$$

$$C_3 = B_1^c \cap B_2^c \cap B_3$$

$$\vdots$$

$$C_n = B_1^c \cap B_2^c \cap \ldots \cap B_{n-1}^c \cap B_n$$

This gives C_i's to be disjoint and $B_i \in \mathcal{S} \implies B_i^c = \bigcup_{j_i=1}^{n_i} S_{ij_i}$ with $S_{ij_i} \in \mathcal{S}$ and are disjoint. Hence,

$$C_1 = B_1 \in \mathcal{S}$$

$$C_2 = B_2 \cap (\cup_{j_1=1}^{n_1} S_{1j_1}) = \bigcup_{j_1=1}^{n_1} (S_{1j_1} \cap B_2)$$

$$C_3 = B_3 \cap (\cup_{j_1=1}^{n_1} S_{1j_1}) \cap (\cup_{j_2=1}^{n_2} S_{2j_2}) = \bigcup_{j_1=1}^{n_1} \bigcup_{j_2=1}^{n_1} S_{1j_1} \cap S_{2j_2} \cap B_3$$

$$\vdots$$

$$C_n = B_n \cap (\cup_{j_1=1}^{n_1} S_{1j_1}) \cap (\cup_{j_2=1}^{n_2} S_{2j_2}) \cap \ldots \cap (\cup_{j_n=1}^{n_n} S_{nj_n})$$

$$= \bigcup_{j_1=1}^{n_1} \bigcup_{j_2=1}^{n_2} \ldots \bigcup_{j_n=1}^{n_n} (B_n \cap S_{1j_1} \cap S_{2j_2} \cap \ldots \cap S_{nj_n})$$

Hence, define :

$$B_{2,j_1} = B_2 \cap S_{1j_1} \text{ for } 1 \leq j_1 \leq n_1$$

And similarly define,

$$B_{3j_1,j_2} = B_3 \cap S_{1j_1} \cap S_{2,j_2}$$

$$\vdots$$

$$B_{n,j_1,\ldots,j_{n-1}} = B_n \cap S_{1,j_1} \ldots S_{n-1,j_{n-1}}$$

Then,

$$C_1 = B_1$$

$$C_2 = \bigcup_{j=1}^{n_2} B_{2j}$$

$$\vdots$$

$$C_n = \bigcup_{j_1=1}^{n_1} \cdots \bigcup_{j_{n-1}=1}^{n_n} B_{n,j_1,\ldots,j_{n-1}}$$

Since C_i's are disjoint elements of $\mathcal{S}$, we have $\{B_1, B_{2,j_1}, B_{3j_1,j_2}, \ldots, B_{n,j_1,\ldots,j_{n-1}}\}$ to be disjoint elements and

$$A = \bigcup_{i=1}^{N} B_i = \bigcup_{i=1}^{N_2} C_i$$

for some N_2 i.e every element of $A \in \mathcal{A}$ is written as the disjoint union of elements from $\mathcal{S}$, which may not be unique.

(Step 2) : Let $A = \bigcup_{i=1}^{N_1} A_i = \bigcup_{i=1}^{N_2} B_i$, $A_i, B_i \in \mathcal{S}$ with A_i's disjoint and B_i's disjoint. Hence,

$$A_k = \bigcup_{i=1}^{N_2} (B_i \cap A_k) \ ; \ B_j = \bigcup_{i=1}^{N_1} (B_j \cap A_i)$$

Hence,

$$m(A_k) = \sum_{i=1}^{N_2} m(B_i \cap A_k)$$

$$\sum_{k=1}^{N_1} m(A_k) = \sum_{k=1}^{N_1} \sum_{i=1}^{N_2} m(B_i \cap A_k)$$

Summing over leads to

$$= \sum_{i=1}^{N_2} m(\cup(B_i \cap A_k)) = \sum_{i=1}^{N_2} m(B_i)$$

Hence $\overline{m}(A) = \sum m(A_k)$ is independent of representation and hence is well defined.

(Step 3) : Let $A = \bigcup_{i=1}^{N} A_i$, $A, A_i \in \mathcal{A}$ with A_i's disjoint. Then,

$$\overline{m}(A) = \sum_{i=1}^{N} \overline{m}(A_i)$$

From step (1),

$$A_i = \bigcup_{j=1}^{N_i} S_{ij}\ ,\ S_{ij} \in \mathcal{S} \text{ with } S_{ij}\text{'s to be disjoint.}$$

$$A = \bigcup_{i=1}^{N} \bigcup_{j=1}^{N_i} S_{ij}$$

Since A_i's are disjoint , S_{ij}'s are disjoint and hence,

$$\overline{m}(A) = \sum_{i=1}^{N} \sum_{j=1}^{N_i} m(S_{ij}) = \sum_{i=1}^{N} \overline{m}(A_i)$$

(Step 4) : Let $A = \bigcup_{i=1}^{\infty} A_i$ with A_i's to be disjoint and $A, A_i \in \mathcal{A}$ and A_i's are disjoint. Let

$$A = \bigcup_{i=1}^{N} S_i, S_i \in \mathcal{S} \implies \overline{m}(A) = \sum_{l=1}^{N} m(S_l)$$

$$A_i = \bigcup_{j=1}^{N_i} S_{ij}\ ;\ S_{ij} \in \mathcal{S}\ ;\ S_{ij}\text{'s are disjoint} \implies \overline{m}(A_i) = \sum_{j=1}^{N_i} m(S_{ij})$$

Then,

$$\bigcup_{i=1}^{N} S_i = A = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{N_i} S_{ij}$$

$$S_l = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{N_i} (S_{ij} \cap S_l) \text{ for all } 1 \le l \le N$$

Since $S_{jk} \cap S_l \in \mathcal{S}$, $(S_{ij} \cap S_l) \cap (S_{pr} \cap S_l) = (S_{ij} \cap Spr) \cap S_l = \phi$ is $(i,j) \neq (p,r)$. Now, by the definition of m, we have

$$m(S_l) = \sum_{i=1}^{\infty} \sum_{j=1}^{N_i} m(S_{ij} \cap S_l)$$

$$\overline{m}(A) = \sum_{l=1}^{N} m(S_l) = \sum_{l=1}^{N} \sum_{i=1}^{\infty} \sum_{j=1}^{N_i} m(S_{ij} \cap S_l)$$

$$S_l \cap A_i = \bigcup_{j=1}^{N_i} (S_l \cap S_{ij}) \implies \overline{m}(S_l \cap A_i) = \sum_{j=1}^{N_i} m(S_l \cap S_{ij})$$

Hence,

$$\overline{m}(A_i) = \sum_{l=1}^{N} \overline{m}(S_l \cap S_{ij}) = \sum_{l=1}^{N} \sum_{j=1}^{N_i} \overline{m}(S_l \cap A_i) = \sum_{l=1}^{N} \sum_{j=1}^{N_i} m(S_l \cap S_{ij})$$

$$\implies \sum_{i=1}^{\infty} \overline{m}(A_i) = \sum_{i=1}^{\infty}\sum_{l=1}^{N}\sum_{j=1}^{N_i} m\big(S_l \cap S_{ij}\big) = \overline{m}(A).$$

□

1.5.2 Construction of measure on X for a given measure on Algebra.

Let $\mathcal{S} \subset \mathbb{P}(X)$ be a semi algebra and m be a measure on S. Let $\mathcal{A}$ be the algebra generated by $\mathcal{S}$ and $\overline{m}$ is the extension of m on $\mathcal{A}$. For $E \subset X$, define

$$m^*(E) = \inf \left\{ \textstyle\sum_{i=1}^{\infty} \overline{m}(A_i) \ ; \ E \subset \bigcup_{i=1}^{\infty} A_i \ , \ A_i \in \mathcal{A} \right\}$$

For $E \subset F$, and if $F = \bigcup_{i=1}^{\infty} A_i$, then $E \subset \bigcup_{i=1}^{\infty} A_i \implies m^*(E) \leq m^*(F)$.
Hence, by the caratheodory's extension theorem, there is a complete $\sigma-$ algebra $\mathcal{F} \subset \mathbb{P}(X)$ of measurable sets.

Theorem 1.53.

$$\mathcal{S} \subset \mathcal{F}$$

Proof. We need to show that if $A \in \mathcal{S}$, then A is measurable and $m^*(A) = \overline{m}(A)$.
Since $A \subset A \in \mathcal{S}$, $m^*(A) \leq \overline{m}(A)$
Fix $\epsilon > 0$ and choose $A_i \subset \mathcal{A}$ such that $A \in \bigcup_{i=1}^{\infty} A_i$ and

$$m^*(A) \geq \sum_{i=1}^{\infty} \overline{m}(A_i) - \epsilon$$

Declare

$$B_i = \begin{Bmatrix} A_i & \text{if } i = 1 \\ A_i \cap A_1^c \cap \ldots \cap A_{i-1}^c & \text{else} \end{Bmatrix}$$

Thus B_i's are disjoint and $\bigcup A_i = \bigcup B_i \implies A \subset \bigcup_{i=1}^{\infty} B_i$.
Also note that $B_i \subset A_i \implies \overline{m}(B_i) \leq \overline{m}(A_i)$.
Note $A \in \mathcal{S}$, $A \subset \bigcup_i B_i \implies A = \bigcup_i (A \cap B_i)$; $A \cap B_i \in \mathcal{A}$ with $(A \cap B_i)$'s are disjoint.
Hence,

$$\overline{m}(A) = \sum_{i=1}^{\infty} \overline{m}(A \cap B_i) \leq \sum_{i=1}^{\infty} \overline{m}(B_i) \leq \sum_{i=1}^{\infty} \overline{m}(A_i)$$

Hence,

$$\overline{m}(A) \leq \sum_{i=1}^{\infty} \overline{m}(A_i) \leq m^*(A) + \epsilon$$

$$\implies m^*(A) = \overline{m}(A)$$

Claim: $A \in \mathcal{S}$ implies A is measurable.
Let $E \subset X$ be any set and let $\epsilon > 0$ such that $E \subset \bigcup_{i=1}^{\infty} A_i$ with $A_i \in \mathcal{A}$.

$$m^*(E) \geq \sum_{i=1}^{\infty} \overline{m}(A_i) - \epsilon$$

Let $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ with $B_i \in \mathcal{A}$ and B_i's are disjoint with $B_i \subset A_i$ implies $\overline{m}(B_i) \leq \overline{m}(A_i)$. Hence,

$$m^*(E) \geq \sum_{i=1}^{\infty} \overline{m}(A_i) - \epsilon \geq \sum_{i=1}^{\infty} \overline{m}(B_i) - \epsilon$$

Now $A \cap E \subset \bigcup_i (A \cap B_i)$ and $A^c \cap E \subset \bigcup_i (A^c \cap B_i)$.
Hence, $B_i = (B_i \cap A) \bigcup (B_i \cap A^c)$ and this implies $\overline{m}(B_i) = \overline{m}(B_i \cap A) + \overline{m}(B_i \cap A^c)$.
Also, by the previous part of the proof, $m^*(B_i) = \overline{m}(B_i)$ and thus,

$$\sum_{i=1}^{\infty} \overline{m}(B_i) \geq \sum_{i=1}^{\infty} \overline{m}(B_i \cap A) + \sum_{i=1}^{\infty} \overline{m}(B_i \cap A^c)$$

Since $E \subset A \subset \bigcup_i (B_i \cap A)$ with $B_i \cap A \in \mathcal{A}$, by the definition of m^*,

$$m^*(E \cap A) \leq \sum_i m^*(B_i \cap A)$$

Similarly

$$m^*(E \cap A^c) \leq \sum_i m^*(B_i \cap A^c)$$

Hence,

$$m^*(E) \geq \sum_{i=1}^{\infty} m^*(B_i) - \epsilon = \sum_{i=1}^{\infty} \overline{m}(B_i \cap A) + \sum_{i=1}^{\infty} \overline{m}(B_i \cap A^c) - \epsilon$$
$$= m^*(E \cap A) + m^*(E \cap A^c) - \epsilon$$

ϵ was arbitrary and hence sending $\epsilon \to 0$ we get the result of the theorem, i.e,

$$m^*(E) \geq m^*(E \cap A) + m^*(E \cap A^c)$$

$$\implies m^*(A) = \overline{m}(A)$$

Hence, $A \in \mathcal{F}$ and $m^* = \overline{m}$ on $\mathcal{A}$.

$\square$

- Let $X \neq \phi$ and $\mathcal{F}_1, \mathcal{F}_2$ be 2 sigma algebras in $\mathbb{P}(X)$. Assume:

- $\mathcal{F}_1 \subset \mathcal{F}_2$
- There are measures m_1 and m_2 on $\mathcal{F}_1$ and $\mathcal{F}_2$ respectively such that $m_1 = m_2$ on $\mathcal{F}_1$.

Let $\overline{\mathcal{F}_1}$ and $\overline{\mathcal{F}_2}$ be the caratheodory completion of $\mathcal{F}_1$ and $\mathcal{F}_2$ with respect to m_1 and m_2 and denote the extenxion of measures to be m_1^* and m_2^* respectively. Then,

- If $A \subset \overline{\mathcal{F}_1}$ and $m_1^*(A) = 0$, then $m_2^*(A) = 0$.
- $m_1^* = m_2^*$ on $\overline{\mathcal{F}_1}$
- $\overline{\mathcal{F}_1} \subset \overline{\mathcal{F}_2}$

Proof. Let $A \subset X$, then, for $j = 1, 2$,

$$m_j^*(A) \equiv \inf\{\sum_{i=1}^{\infty} m_j(B_i) \ ; \ A \subset \bigcup_{i=1}^{\infty} B_i \ , \ B_i \in \mathcal{F}_j\}$$

Then $m_2^*(A) \leq m_1^*(A)$. Hence, $m_1^*(A) = 0 \implies m_2^*(A) = 0$. But $A \in \mathcal{F}_1$ implies $A \in \mathcal{F}_2$. This shows that $\overline{\mathcal{F}_1} \subset \overline{\mathcal{F}_2}$.

$\square$

Lemma 1.54. *Some applications include*

- ***Lebesgue measure on*** $\mathbb{R}$ ***:***
  Let

$$\mathcal{S} = \{(a, b] \ ; \ a, b \in \mathbb{R}\} \bigcup \{(a, \infty) \ ; \ a \in \mathbb{R}\} \bigcup \{(-\infty, b) \ ; \ b \in \mathbb{R} \ \}$$

$$m((a, b]) = b - a$$

Claim : $\mathcal{S}$ *is a semi algebra and* m *is a measure.*

Proof.

$$(a_1, b_1] \cap (a_2, b_2] = \left\{ \begin{array}{cc} \phi & \text{if } b_1 \leq a_2 \text{ or } b_2 \leq a_1 \\ (max(a_1, a_2), min(b_1, b_2)] & \text{otherwise} \end{array} \right\} \in \mathcal{S}$$

Now,

$$\mathbb{R} - (a, b] = (-\infty, a] \bigcup (b, \infty) \in \mathcal{S}$$

Let $[a, b] = \bigcup_{i=1}^{\infty} (a_i, b_i]$ with $(a_i, b_i]$ disjoint. Let $\epsilon > 0$ such that $a + \epsilon < b$. Then

$$[a + \epsilon, b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i + \frac{\epsilon}{2^i})$$

By compactness, $\exists\, n(\epsilon) \equiv N \in \mathbb{N}$ such that

$$[a+\epsilon, b] \subset \bigcup_{i=1}^{N} (a_i, b_i + \frac{\epsilon}{2^i})$$

Hence,

$$b - a - \epsilon \leq \sum_{i=1}^{N} (b_i - a_i) + \epsilon \sum_{i=1}^{N} \frac{1}{2^i}$$

$$\leq \sum_{i=1}^{N} (b_i - a_i) + \epsilon$$

$$b - a \leq \sum_{i=1}^{\infty} (b_i - a_i) - 2\epsilon$$

$\epsilon > 0$ was arbitrary and hence,

$$m((a,b]) = b - a \leq \sum_{i=1}^{\infty} m((a_i, b_i])$$

Now, $\forall N > 0$,

$$\begin{Bmatrix} \bigcup_{i=1}^{N} (a_i, b_i] \subset (a,b] \\ \{(a_i, b_i]\} \text{ are disjoint.} \end{Bmatrix}$$

Choose $\epsilon > 0$ such that $a_i < b_i - \frac{\epsilon}{2^i}$ $\forall 1 \leq i \leq N$. Then

$$\bigcup_{i=1}^{N} (a_i, b_i - \frac{\epsilon}{2^i}) \subset \bigcup [a_i, b_i] \subset (a,b]$$

And $(a_i, b_i - \frac{\epsilon}{2^i})$'s are mutually disjoint. Hence,

$$\sum_{i=1}^{N} (b_i - a_i) - \epsilon \sum_{i=1}^{N} \frac{1}{2^i} \leq b - a$$

let $\epsilon \to 0$ to get

$$\sum_{i=1}^{N} (b_i - a_i) \leq b - a$$

Let $N \to \infty$ to get

$$\sum_{i=1}^{\infty} m(a_i, b_i] \leq \sum_{i=1}^{\infty} (b_i - a_i) \leq b - a = m(a,b]$$

Hence,

$$m(a,b] = \sum_{i=1}^{\infty} m(a_i, b_i]$$

Thus , m is a measure on $\mathcal{S}$. Hence, by the theorem, there is a complete $\sigma-$ algebra $\mathcal{F} \subset \mathbb{P}(\mathbb{R})$ and a measure m^* on $\mathcal{F}$ such that

- $\mathcal{S} \in \mathcal{F}$
- $m^* = m$ on $\mathcal{S}$

$\sigma(\mathcal{S}) \subset \mathcal{F}$ and $\sigma(\mathcal{S})$ is Borel $\sigma-$ algebra on $\mathbb{R}$. $\mathcal{F}$ is called the lebesgue measure space with $m^* = dx$ to be the lebesgue measure.

□

- ***Lebesgue measure on*** $\mathbb{R}^n$ ***:***

Let

$\mathcal{S} = \left\{\prod_{i=1}^n (a_i, b_i] \text{ ; } a_\infty = -\infty \text{ , } b_i = \infty \textit{ are included and if } b_i = \infty, \textit{ then } (a_i, b_i] = (a_i, \infty)\right\}$

Claim : $\mathcal{S}$ *is a semi algebra.*
Proof of the claim :

$$\left(\prod_{i=1}^n (a_i, b_i]\right) \cap \left(\prod_{i=1}^n (c_i, d_i]\right) = \prod_{i=1}^n ((a_i, b_i] \cap (c_i, d_i])$$

$$= \prod_{i=1}^n (\alpha_i, \beta_i] \in \mathcal{S}$$

Let $\mathbb{P}(n)$ *denote the power set of* $\{1, 2 \ldots, n\}$. *Then, by induction,*

$$\left(\prod_{i=1}^n (a_i, b_i]\right)^c = \prod_{J \in \mathbb{P}(n)} D_J$$

where, D_J *is the disjoint collection of sets in of the form : Let* $J \equiv (j_1, j_2, \ldots, j_k)$, *then*

$$D_J = A_1 \times A_2 \times \ldots \times A_n, \textit{ where}$$

$$A_i = (a_i, b_i] \textit{ if } i \notin J$$

$$A_i = (a_i, b_i]^c \equiv (-\infty, a_i] \cup (b_i, \infty) \textit{ if } i \in J$$

Hence, $\mathcal{S}$ *is a semi algebra.*
Claim : m *is a measure on* $\mathcal{S}$.
Proof of the claim : For

$$\prod_{i=1}^n (a_i, b_i] = \bigcup_{j=1}^\infty \prod_{i=1}^n (c_i^j, d_i^j]$$

they are disjoint and thus

$$\prod_{i=1}^{n} \mathcal{X}_{(a_i,b_i]}(x_i) = \sum_{j=1}^{\infty} \prod_{i=1}^{n} \mathcal{X}_{(c_i^j,d_i^j]}(x_i)$$

Now, in each variable, dx_i is the lebesgue measure.
Apply DCT to get

$$\prod_{i=1}^{n} (b_i - a_i) = \sum_{j=1}^{n} \prod_{i=1}^{n} (d_i^j - c_i^j)$$

Hence, there is a complete $\sigma-$ algebra $\mathcal{F}_0$ and a measure $\overline{m}$ denoted by $dx \equiv dx_1.dx_2 \ldots dx_n$ such that $\mathcal{S} \subset \mathcal{F}_0$ and $\overline{m} = m$ on $\mathcal{S}$.
Here $\overline{m} = dx$ is called the n-dimensional lebesgue measure.

□

- *Let $(X, \mathbb{A}, \mu)$ and $(Y, \mathbb{B}, v)$ be 2 measure spaces and $Z = X \times Y$. Let*

$$\mathcal{S} = \{A \times B \ ; \ A \in \mathbb{A} \ , \ B \in \mathbb{B}\} \subset \mathbb{P}(Z)$$

$$\lambda(A \times B) = \mu(A).v(B)$$

Then $\mathcal{S}$ is a semi algebra and λ is a measure on $\mathcal{S}$.

Proof.

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \times B_2)$$

$$(A \times B)^c = (A^c \times B) \cup (A \times B^c) \cup (A^c \times B^c)$$

Thus $\mathcal{S}$ is a semialgebra. Let

$$A \times B = \bigcup_{i=1}^{\infty} (A_i \times B_i)$$

for $A_i, A \in \mathbb{A}$, $B_i, B \in \mathbb{B}$ and $A_i \times B_i$'s are disjoint.
Then, for $x \in X, y \in Y$, we have,

$$\mathcal{X}_A(x)\mathcal{X}_B(y) = \mathcal{X}_{A\times B}(x,y) = \sum_{i=1}^{\infty} \mathcal{X}_{A_i \times B_i}(x,y)$$

$$= \sum_{i=1}^{\infty} \mathcal{X}_{A_i}(x).\mathcal{X}_{B_i}(y)$$

Hence, by MCT, we have $\forall y \in Y$, we have

$$\mu(A)\mathcal{X}_B(y) = \sum_{i=1}^{\infty} \mu(A_i)\mathcal{X}_{B_i}(y)$$

$$\implies \mu(A)v(B) = \sum_{i=1}^{\infty} \mu(A_i)v(B_i)$$

$$\implies \lambda(A \times B) = \sum_{i=1}^{\infty} \lambda(A_i \times B_i)$$

Hence there is a complete $\sigma-$ algebra $\mathcal{F} \subset \mathbb{P}(Z)$ and a measure λ^* on $\mathcal{F}$ such that

- $\mathcal{S} \subset \mathcal{F}$
- $\lambda^* = \lambda$ on $\mathcal{S}$

λ^* is called the **product measure** and sometimes denoted by $\mu \times v$.

□

Let $\{(X_i, \mathcal{F}_i, m_i)\}_{i \in I}$ be a family of probability spaces, i.e, $m_i(X_i) = 1$ for all $i \in I$. Let

$$X = \prod_{i \in I} X_i := \{(x_i)_{i \in I} \ ; \ x_i \in X_i\}$$

Let

$$\mathcal{S} = \left\{ \prod_{i \in I} A_i \ ; \ A_i \in \mathcal{F}_i, A_i = X_i \text{ for all but finitely many } i \right\}$$

$$\lambda \left(\prod_{i \in I} A_i \right) := \prod_{i \in I} m_i(A_i)$$

Claim : S is a semi algebra and λ is a measure.
Proof of the claim :

$$\prod_{i \in I} A_i \cap \prod_{i \in I} B_i = \prod_{i \in I} (A_i \cap B_i) \in \mathcal{S} \ ; \text{ for } \prod_{i \in I} A_i, \prod_{i \in I} B_i \in \mathcal{S}$$

Since $A_i \cap B_i = X_i$ for all, but finitely many i. Let $\alpha = \prod_{i \in I} A_i$, let

$$J_0 := J_0(\alpha) = \{i \in I \ ; \ A_i \neq X_i\}$$

The clearly $Card(J_0) < \infty$. Hence, for $J \subset J_0$, define

$$\beta_J = \prod_{i \in I} C_i^J$$

where

$$C_i^J = \begin{Bmatrix} X_i & \text{if } i \notin J \\ A_i^c & \text{if } i \in J \end{Bmatrix}$$

Thus $\beta_J \cap \beta_L = \phi$, if $J \neq L$ and

$$\alpha = \bigcup_{J \subset J_0} \beta_J = \bigcup_{J \subset J_0} \prod_{i \in J} C_i^J$$

And hence $\mathcal{S}$ is a semi algebra.
Define $m : \mathcal{S} \to [0,1]$ as

$$m\left(\prod A_i\right) = \prod m_i(A_i) = \prod_{i\in I; A_i \neq X_i} m_i(A_i)$$

Claim : m defined as above, is a measure on the semi algebra.
Proof of the claim : Suppose

$$\prod_{i\in I} A_i = \bigcup_{j=1}^{\infty}\left(\prod_{i\in I} C_i^j\right) \text{ , for } \prod_{i\in I} C_i^j \cap \prod_{i\in I} C_i^k = \phi \text{ , if } j \neq k$$

Then, for $x \equiv (x_i)_{i\in I}$,

$$\prod_{i\in J_0} \mathcal{X}_{A_i}(x_i) = \mathcal{X}_{\prod A_i}(x) = \sum_{j=1}^{\infty} \mathcal{X}_{\prod C_i^j}(x)$$

$$= \sum_{j=1}^{\infty} \prod_{i\in J_j} \mathcal{X}_{C_i^j}(x_i)$$

where

$$J_0 = \{i \in I \ ; \ A_i \neq X_i\}$$

$$J_j = \{i \in I \ ; \ C_i^j \neq X_i\}$$

Then, clearly $Card(J_0) < \infty$, $Card(J_j) < \infty$. By DCT,

$$\prod_{i\in I} \mathcal{X}_{A_i}(x_i) = \sum_{j=1}^{\infty} \prod_{i\in J_j} \mathcal{X}_{C_i^j}(x_i)$$

$$\prod_{i\in J_0} m_i(A_i) = m\left(\prod_{i\in I} A_i\right) = \sum_{j=1}^{\infty}\left(\prod_{i\in J_j} m_i(C_i^j)\right) = \sum_{j=1}^{\infty}\left(\prod_{i\in I} C_i^j\right)$$

Hence, m is a measure on $\mathcal{S}$. Thus there is a complete $\sigma-$algebra $\mathcal{F}_0 \supset \mathcal{S}$ and a measure $\overline{m}$ on $\mathcal{F}_0$ such that $\overline{m} = m$ on $\mathcal{S}$. Here $\overline{m}$ is called the product measure.

$\square$

Lemma : *Let X be a non-empty set and $g : X \to [0,\infty]$ be a function. Define A_k For $k \in \mathbb{N} \cup \{0\}$ as ,*

$$(k = 0) \ , \ A_0 := \phi$$

$$(k = 1) \ , \ A_1 := \{x \ ; \ g(x) \geq 1\}$$

$$(\textit{ For } k \geq 2) \ , \quad A_k := \{x \ ; \ g(x) \geq \frac{1}{k} + \sum_{i=1}^{k-1} \frac{1}{i} \mathcal{X}_{A_i}(x)\}$$

Then,

$$g(x) \equiv \sum_{i=1}^{\infty} \frac{1}{i} \mathcal{X}_{A_i}(x)$$

Furthermore, if g is a measurable function, then A_i's are measurable sets.

Proof. Suppose that $g(x) = \infty$, then $x \in A_i$, for all i. Hence,

$$\sum_{i=1}^{\infty} \frac{1}{i} \mathcal{X}_{A_i}(x) = \sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

Let $0 \leq g(x) < \infty$. Since, for any $k > 0$, $\sum_{i=k}^{\infty} \frac{1}{i} = \infty$, there are infinite sets $K, L \subset \mathbb{N}$ such that $K \cap L = \phi$ and $\mathbb{N} = K \cup L$.
$x \notin A_i$ for $i \in K$ and $a \in A_l$, for all $l \in L$, define $S : K \to L \cup \{0\}$ as

$$S(j) = 0 \text{ if } x \notin A_i \ , \ \forall i < j$$

$$S(j) = \max\{i \ ; \ i < j \ , \ x \in A_i\}$$

Hence, for all $j \in K$, $x \notin A_j$ and $x \in A_{S(j)}$, if $S(j) \neq \{0\}$.

$$\Longrightarrow \ \frac{1}{S(j)} + \sum_{i=1}^{S(j)-1} \frac{1}{i} \mathcal{X}_{A_i}(x) \leq g(x) \leq \frac{1}{j} + \sum_{i=1}^{j-1} \frac{1}{i} \mathcal{X}_{A_i}(x)$$

$$(i.e) \ \sum_{i<j} \frac{1}{i} \mathcal{X}_{A_i}(x) \leq g(x) \leq \frac{1}{j} + \sum_{\{i<j \ ; \ i \notin K\}} \frac{1}{i} \mathcal{X}_{A_i}(x)$$

Since K is an infinite set, letting $j \to \infty$, we get

$$g(x) = \sum_{\{i=1 \ ; \ i \notin K\}}^{\infty} \frac{1}{i} \mathcal{X}_{A_i}(x) = \sum_{i=1}^{\infty} \frac{1}{i} \mathcal{X}_{A_i}(x)$$

□

Remark : In more generality, if $r_i > 0$, $\sum_{i=1}^{\infty} r_i = \infty$, define

$$A_0 := \phi$$

$$A_1 := \{x \in X \ ; \ f(x) \geq r_1\}$$

$$\text{For } k \geq 2,\ A_k := \{x \in X \ ;\ f(x) \geq r_k + \sum_{i=1}^{k-1} r_i \mathcal{X}_{A_i}(x)\}$$

Then,

$$f(x) \equiv \sum_{i=1}^{\infty} r_i \mathcal{X}_{A_i}(x)$$

Proof of the remark follows the same proof as in the previous lemma with $r_i = \frac{1}{i}$. $\square$

1.5.3 Geometry and measure theory on $\mathbb{R}$

Denote m^* to be the outer lebesgue measure on $\mathbb{R}$, that is

$$m^*(A) = \inf \left\{ \textstyle\sum_{i=1}^{\infty} l(I_i) \ ;\ A \subset \bigcup_{i=1}^{\infty} I_i \right\}$$

$$I_i := (a_i, b_i)\ ,\ l(I_i) = b_i - a_i$$

Lemma 1.55. ***Vitali** :*
Let E be a set of finite outer measure and $\mathbb{I}$ be a collection of intervals which cover E in the sense of vitali i.e given $\epsilon > 0$ and $x \in E, \exists I \in \mathbb{I}$ such that $x \in I$ and $l(I) < \epsilon$. Then given $\epsilon > 0$ there is a finite set of disjoint intervals $\{I_1, \ldots, I_n\}$ of intervals in $\mathbb{I}$ such that

$$m^*\left[E - \bigcup_{i=1}^{n} I_i\right] < \epsilon$$

Proof. Since, the endpoints of an interval is of measure zero, W.L.O.G we can assume the intervals in $\mathbb{I}$ are closed. Also, since E is a set of finite outer measure, we can assume that $E \subset O$, for some open set O of finite measure. The collection of intervals from $\mathbb{I}$ which are contained in O, also form a vitali cover. Consider this new set to be $\mathbb{I}$. Let $I \in \mathbb{I}$, then $I \subset O$ and thus $l(I) \leq m(O) < \infty \implies \sup\{l(I) \ ;\ I \in \mathbb{I}\} \leq m(O) < \infty$.
Algorithm for choosing the disjoint intervals:
Let I_1 be any interval of choice from $\mathbb{I}$.
If $E \subset I_1$, then nothing is to be done. Else, let

$$\mathcal{F}_2 := \{I \in \mathbb{I} \ ;\ I \cap I_1 = \phi\}$$

$$k_2 := \sup\{l(I) \ ;\ I \in \mathcal{F}_2\}$$

Now, choose $I_2 \in \mathcal{F}_2$ such that $l(I_2) \geq \frac{k_2}{2}$.
By induction, suppose that $I_1, I_2, \dots I_{j-1}$ are chosen, then define

$$\mathcal{F}_j := \{I \in \mathbb{I} \; ; \; I \cap I_k = \phi, \; \forall k \leq j-1\}$$

$$k_j := \sup\{l(I) \; ; \; I \in \mathcal{F}_j\}$$

Choose $I_j \in \mathcal{F}_j$ such that $l(I_j) \geq \frac{k_j}{2}$.
This way, we generate a sequence of intervals $\{I_n\}$ and $k_n < m(O) < \infty, \forall n$.
Now $\cup I_n \subset O$ implies $\sum l(I_n) \leq m(O) < \infty$. Hence, given $\epsilon > 0, \exists N$ such that $\sum_{N+1}^{\infty} l(I_n) < \frac{\epsilon}{5}$.
Declare

$$R = E - \bigcup_{n=1}^{N} I_n$$

Note that $k_n \to 0$ as it is a tail of a convergent series.

Claim : $m^* R < \epsilon$ which is the exact conclusion that is needed for the theorem.
Proof of the claim : Let $x \in R$ implies $x \in E$ and $x \notin \cup_{i=1}^{N} I_i$, which is a closed set by assumption and hence, $d(x, \cup_{i=1}^{N} I_i) > 0$.
Since the cover is vitali, there is an $I \in \mathbb{I}$ such that $x \in I$ and $I \cap I_j = \phi$, for all $j = 1, 2, \dots, N$.
Suppose that $I \cap I_j = \phi$, $\forall j$, then $l(I) \leq k_j$ for all j and as $k_n \to_{n\to\infty} 0$, $l(I) = 0$ which is a contradiction.
Hence, let $n \geq N+1$ be the smallest integer such that $I \cap I_j = \phi$ for all $j \leq n-1$ and $I \cap I_n \neq \phi$. Hence, $I \in \mathcal{F}_n$ and thus $l(I) \leq k_n \leq 2l(I_n)$. Let $z \in I \cap I_n$ and m be the midpoint of I_n, then for $a \in I$,

$$dist(a, m) \leq dist(a, z) + dist(z, m) \leq l(I) + \frac{l(I_n)}{2} \leq 2l(I_n) + \frac{l(I_n)}{2} = \frac{5}{2} l(I_n)$$

Let J_n be the interval around m , but with 5 times the radius of I_n. Hence, J_n covers I, I_n, x, m. This implies $x \in J_n$. Hence $R \subset \bigcup_{n=N+1}^{\infty} J_n$.

$$m^* R \leq \sum_{N+1}^{\infty} l(J_n) \leq 5 \sum_{N+1}^{\infty} l(I_n) \leq 5 . \frac{\epsilon}{5} = \epsilon$$

$\square$

Theorem 1.56. *For $f : [a, b] \to \mathbb{R}$ to be increasing monotone function,*

- $f' = \frac{df}{dx}$ *exists for almost every $x \in [a, b]$.*
- f' *is measurable.*

- $\int_a^b f'(x)dx \leq f(b) - f(a)$.

Proof. Define 4 quantities as follows :

$$D^+f(x) = \limsup_{h \to 0+} \frac{f(x+h) - f(x)}{h}$$

$$D^-f(x) = \limsup_{h \to 0-} \frac{f(x+h) - f(x)}{h}$$

$$D_+f(x) = \liminf_{h \to 0+} \frac{f(x+h) - f(x)}{h}$$

$$D_-f(x) = \liminf_{h \to 0-} \frac{f(x+h) - f(x)}{h}$$

And f is said to be differentiable at x if all of the 4 quantities are same. Clearly, $D^+f(x) \geq D_+f(x)$ and $D^-f(x) \geq D_-f(x)$

The idea is to show that the set where the 2 derivatives are unequal is of measure zero, i.e we want the set $\{x : D^+f(x) > D_-f(x)\}$ and $\{x : D_+f(x) > D^-f(x)\}$ to have measure zero. Then, outside this set, all the derivatives are equal.

Consider the set $E = \{x : D^+f(x) > D_-f(x)\}$. The other case is similar.

$$E = \bigcup_{u,v \in \mathbb{Q}} E_{u,v} := \bigcup_{u,v \in \mathbb{Q}} \Big\{ x : D^+f(x) > u > v > D_-f(x) \Big\}$$

where $\mathbb{Q}$ is the set of rationals. It is good enough to show that $m^*(E_{u,v}) = 0$. Since $[a, b]$ has finite measure, let $s = m^*(E_{u,v})$. Note, it is not yet known if $E_{u,v}$ is measurable, hence the outer measure. Now, get an open set O such that $E_{u,v} \subset O$ and $m(O) < s + \epsilon$, by the outer regularity.

Now, $x \in E_{u,v} \subset O \implies \exists h' > 0$ such that $B(x, h') \subset O$. Modify with the condition with $D_-(f) < v$ to get, for small $h < h'$, $f(x) - f(x-h) < v.h$. Thus all, $I_x = [x-h, x]$ forms vitali cover for $E_{u,v}$.

By the previous theorem , there are finite intervals $I_1 \ldots I_N$ such that $m^*(E_{u,v} - \bigcup_{i=1}^N I_i) < \epsilon$ and as a consequence,

$$\sum_{i=1}^{N} \big[f(x_i) - f(x_i - h_i) \big] \leq \sum_{i=1}^{N} v.h_i \leq v.m(O) < v(s + \epsilon)$$

Let the interior of these intervals cover a set A along with $A \cap E_{u,v} \subset A$. Since,

$$A = \bigcup_i I_i^\circ$$

$$E_{u,v} = (E_{u,v} \cap A) \cap (E_{u,v} - A)$$

we have

$$s = m^*(E_{u,v}) \leq m^*(E_{u,v} \cap A) + m^*(E_{u,v} - A) \leq m^*(E_{u,v} \cap A) + \epsilon$$

Hence,

$$m^*(E_{u,v} \cap A) \geq s - \epsilon$$

Repeating the above logic of generating vitali cover for $A \cap E_{u,v}$ and looking at the corresponding sum of images of f, with respect to $D^+(f) > u$, we get $J_1, \dots J_M$ which covers a set B of $A \cap E_{u,v}$ with the properties :

- $m^*(A \cap E_{u,v} - \bigcup_{i=1}^M J_i) < \epsilon$.
- J_i's are disjoint and J_i is of the from $[y_i, y_i + k_i]$.
- Given i, $\exists l$ such that $J_l \subset I_i^\circ$.
- $[f(y_i + k_i) - f(y_i)] \geq u k_i$.

Now,

$$E_{u,v} \cap A = \left(E_{u,v} \cap A \cap (\bigcup_{j=1}^M J_j)\right) \bigcup \left(E_{u,v} \cap A - (\bigcup_{j=1}^M J_j)\right)$$

Hence,

$$s - \epsilon \leq m^*(E_{u,v} \cap A) \leq m^*\left(E_{u,v} \cap A \cap (\bigcup_{j=1}^M J_j)\right) + \epsilon$$

$$\implies m^*\left(E_{u,v} \cap A \cap (\bigcup_{j=1}^M J_j)\right) \geq s - 2\epsilon$$

Now,

$$\sum_{i=1}^M \left[f(y_i + k_i) - f(y_i)\right] > u \sum_{i=1}^M k_i = u \sum_{j=1}^M l(J_j)$$

$$= u\, m\left(\bigcup_{j=1}^M J_j\right)$$

$$\geq u\, m^*\left(E_{u,v} \cap A \cap (\bigcup_{j=1}^M J_j)\right)$$

$$> u(s - 2\epsilon)$$

Summing over those i for which $J_i \subset I_n$,

$$f(x_n) - f(x_n - h_n) \geq \sum_i f(y_i + k_i) - f(y_i)$$

$$\sum_{n=1}^{N} f(x_n) - f(x_n - h_n) \geq \sum_i f(y_i + k_i) - f(y_i)$$

The above is got with the usage of f being non-decreasing function.
Hence, $v(s + \epsilon) > u(s - 2\epsilon)$.
Now $u > v \implies s = 0$ which also shows that $E_{u,v}$ is measurable.
This gives that

$$h(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

is defined almost everywhere and f is differentiable when h is finite.

Set $h_n(x) = n[f(x + 1/n) - f(x)]$ with declaration of $f(x) = f(b)$ for $x \geq b$.
Now, $h_n \to f$ a.e and hence h is measurable. Thus,

$$\int_a^b f'(x)dx \leq \liminf_{k \to \infty} \int_a^b h_n(x)dx$$

$$= \lim_{n \to \infty} n\left[\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} f(x)dx - \int_a^b f(x)dx\right]$$

$$= \lim_{n \to \infty} n\left[\int_b^{b+\frac{1}{n}} f(x)dx - \int_a^{a+\frac{1}{n}} f(x)dx\right]$$

$$= f(b) - f(a+)$$

where $f(a+) := \lim_{x \downarrow a} f(x)$. Since $f(a) \leq f(a+)$, the theorem is proved.

$\square$

Definition 1.57.

- For $f : \mathbb{R} \to \mathbb{R}$ define $f^+ = max(f, 0)$ and $-f^- = min(f, 0)$.
- For $f : [a, b] \to \mathbb{R}$ and partition $a = x_0 \leq x_1 \leq \ldots x_{n-1} \leq x_n = b$ of $[a, b]$ define

$$p = \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})]^+$$

$$n = \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})]^-$$

$$t = n + p = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

- Define P, N, T to be *sup* p , *sup* n and *sup* t, supremum taken over all partitions of $[a, b]$ respectively.
- Sometimes, the above P, N, T are also denoted by P_a^b, N_a^b, T_a^b, referring the interval $[a, b]$.
- $f : [a, b] \to \mathbb{R}$ is said to be bounded variate over $[a, b]$ if T is finite.

Remark 1.58. For notations as above, $t = p + n$ and $P, N \leq T \leq P + N$

Lemma 1.59. *If f is bounded variate over $[a, b]$, then $T = P+N$ and $f(b)-f(a) = P-N$.*

Proof. For any subdivision of $[a, b]$,

$$p = n + f(b) - f(a) \leq N + f(b) - f(a)$$

and hence,

$$P \leq N + f(b) - f(a)$$

Reversing p and n, we get $P - N = f(b) - f(a)$. Hence,

$$T \geq p + n = p + p - [f(b) - f(a)] = 2p + N - P$$

And so,

$$T \geq 2P + N - P = P + N$$

Along with $T \leq P + N$, we get $T = P + N$

□

Theorem 1.60. *$f : [a, b] \to \mathbb{R}$ is bounded variate iff f is a difference of 2 monotone functions.*

Proof. Define P_a^x to be the P for the interval $[a, x]$ and similarly define, N_a^x and T_a^x. By the supremum property, they are increasing real valued. By the previous lemma, $f(x) = P_a^x - N_a^x + f(a)$ which is the one way result needed.
Conversely, if $f = g - h$ on $[a, b]$ with g, h be increasing, then for any subdivision,

$$\sum_i \left|f(x_i)-f(x_{i-1})\right| \leq \sum_i \left[g(x_i)-g(x_{i-1})\right]+\sum_i \left[h(x_i)-h(x_{i-1})\right] = g(b)-g(a)+h(b)-h(a)$$

and hence T is bounded.

□

Corollary 1.61. *If f, for $f : [a,b] \to \mathbb{R}$ is bounded variate, then f is differentiable almost everywhere in $[a,b]$.*

Proof. Monotone functions are differentiable almost everywhere by the earlier theorem and sum of 2 differenctiable functions is differentiable.

□

Lemma 1.62. *For f to be non negative integrable function over a set E. Then given $\epsilon > 0$, there is a $\delta > 0$ such that*

$$\textit{for every } A \subset E,\ m(A) < \delta \implies \int_A f < \epsilon$$

Proof. If f is bounded, say image of f is in $[0, M]$, then, $\int_A f < M\ m(A)$. Choosing δ be $\frac{\epsilon}{M}$ will do the job.
If f is not bounded, Truncate f as sequence with $f_n(x) = f(x)$ if $f(x) \leq n$ and $f_n(x) = n$ elsewhere. Then f_n converges to f pointwise monotonically. Hence by MCT,
$\int_E f - \int_E f_N < \epsilon/2$. Letting $\delta < \frac{\epsilon}{2N}$, we get

$$\int_A f = \int_A (f - f_N) + \int_A f_N < \frac{\epsilon}{2} + N\ m(A) < \epsilon$$

for all A with $m(A) < \delta$.

□

Lemma 1.63. *If f is integrable on $[a,b]$, then the function F defined as*

$$F(x) = \int_a^x f(t)dt$$

is a continuous function of bounded variation on $[a,b]$.

Proof. By the previous lemma, F is continuous.
It is just left to show that F is bounded variate.
For this consider any partition of $[a,b]$ by $a = x_0 < x_1 < \cdots < x_n = b$ then,

$$\sum_{i=1}^{n} \left|F(x_i) - F(x_{i-1})\right| \leq \sum_i \int_{x_i}^{x_{i-1}} |f(t)|dt = \int_a^b |f(t)|dt < \infty$$

This is true for all partitions of $[a,b]$ and T is thus bounded by $\int_a^b |f(t)|dt$.

□

Lemma 1.64. *If f is integrable on $[a,b]$ and*

$$\int_a^x f(t)dt = 0$$

for all $x \in [a,b]$, then $f(t) = 0$ a.e in $[a,b]$.

Proof. (Proof by contradiction.) Suppose f is not 0 a.e, then assume W.L.O.G $f > 0$ on E which has a strict positive measure. By the inner regularity, there is a compact set $K \subset E$ such that $\int_K f > 0$. Now $K^c = \cup(a_i, b_i)$ disjoint intervals. By hypothesis, $\int_{a_i}^{b_i} f = \int_a^{b_i} f - \int_a^{a_i} f = 0$. Let $U = (a,b) - K$. Then

$$0 = \int_a^b f = \int_K f + \int_U f$$

$$\implies \int_U f = -\int_K f$$

Now,

$$\int_K f \neq 0 \implies \int_U f = \sum_i \int_{a_i}^{b_i} f \neq 0$$

But, for all i,

$$\int_{a_i}^{b_i} f = 0 \implies \int_U f = 0$$

This is a clear contradiction.

□

Lemma 1.65. *If f is bounded and measurable on $[a,b]$ and*

$$F(x) = \int_a^x f(t)dt + F(a)$$

then, $F'(x) = f(x)$ for a.e $x \in [a,b]$.

Proof. By the lemma 1.63, F is bounded variate on $[a,b]$ and hence f' exists for almost all $x \in [a,b]$.
Declare

$$f_n(x) = \frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)dt \text{ ; for } h = 1/n$$

Now, $|f| < M \implies |f_n| \leq M$. And Bounded Convergence Theorem along with $f_n \to F'$ for a.e x implies that for any $c \in [a,b]$,

$$\int_a^c F'(x)dx = \lim_{n\to\infty} \int_a^c f_n(x)dx = \lim_{h\to 0} \frac{1}{h}\int_a^c [F(x+h) - F(x)]dx$$

$$= \lim \left[\frac{1}{h}\int_c^{c+h} F(x)dx - \frac{1}{h}\int_a^{a+h} F(x)dx\right] = F(c) - F(a) = \int_a^c f(x)dx$$

The above uses that F is continuous. So, $\int_a^c [F'(x) - f(x)]dx = 0, \quad \forall c \in [a, b]$.
By the previous lemmas, we have $F'(x) = f(x)$ a.e.

□

Lemma 1.66. *Let f be an integrable function on $[a, b]$ and if*

$$F(x) = F(a) + \int_a^x f(t)dt$$

then, $F'(x) = f(x)$ for a.e $x \in [a, b]$.

Proof. W.L.O.G assume $f \geq 0$. Note that F is an increasing monotone function. Truncate f as sequence with $f_n(x) = f(x)$ if $f(x) \leq n$ and $f_n(x) = n$ elsewhere. Then $f - f_n \geq 0$ and so, $h_n(x) = \int_a^x (f - f_n)$ is an increasing function. Hence, it has derivative a.e. Note that this derivative will be positive. Now by the previous lemma,

$$\frac{d}{dx} \int_a^x f_n = f_n(x) \text{ a.e}$$

and hence,

$$F'(x) = \frac{d}{dx} h_n + \frac{d}{dx} \int_a^x f_n \geq f_n(x) \text{ a.e}$$

Now n was arbitrary. This tells that

$$F'(x) \geq f(x) \text{ a.e}$$

Also,

$$\int_a^b F' \geq \int_a^b f = F(b) - F(a)$$

Since F is continuous and increasing, by the theorem 1.56,

$$\int_a^b F'(x)dx \leq F(b) - F(a)$$

$$\implies \int_a^b F' = F(b) - F(a) = \int_a^b f \quad ; \quad \int_a^b [F' - f] = 0$$

Since, $F' - f \geq 0$, it tells that $F' - f = 0$ a.e and so $F'(x) = f(x)$ a.e

□

Definition 1.67. $f : [a, b] \to \mathbb{R}$ is said to be absolutely continuous, if given $\epsilon > 0, \exists \delta > 0$ such that for every collection of finite disjoint intervals $\{(x_i, x_i^{'})\}_{i=1}^n$ with

$$\sum_{i=1}^n |x_i - x_i^{'}| < \delta \implies \sum_{i=1}^n |f(x_i) - f(x_i^{'})| < \epsilon$$

Lemma 1.68. *If f is absolutely continuous on $[a, b]$, then it is of bounded variation on $[a, b]$.*

Proof. Fix $\epsilon = 1$. By the definition of absolute continuity, there exists δ. Now for any subdivision of $[a, b]$, we can split the division further into L set of intervals with each of length less than δ ; where L is the largest natural number less than $\frac{\delta+b-a}{\delta}$.
Therefore, t is always less than L for any subdivision. Hence, f is bounded variate.

□

Corollary 1.69. *If f is absolutely continuous, then f' exists almost everywhere*

Proof. By the previous lemma, absolute continuous functions are bounded variate and hence sum of 2 monotone functions and hence differentiable almost everywhere.

□

Lemma 1.70. *If f is absolutely continuous on $[a, b]$ and it is given that $f'(x) = 0$ a.e, then f is identically a constant function.*

Proof. The main idea is to show $f(a) = f(c)\ \forall c \in [a, b]$. By the hypothesis, let $E \subset (a, c)$ be the set of measure $c - a$ in which $f' = 0$. Let ϵ, η be any 2 positive real numbers. Let δ be the corresponding one for ϵ in the definition of absolute continuity. For each $x \in E$, there is a small interval $[x, x + h] \subset [a, c]$ such that

$$|f(x + h) - f(x)| < \eta h$$

Now, these $\{[x, x + h]\}$ form a vitali cover for E. By the vitali lemma, we can get $\{[x_k, y_k]\}$ such that they are disjoint and cover whole of E except for a set of measure less than δ. Label the intervals such that $x_k \leq x_{k+1}$. We have

$$y_0 = a \leq x_1 < y_1 \leq x_2 < y_2 \leq \cdots \leq y_n \leq c = x_{n+1}$$

and

$$\sum_{k=0}^{n} |x_{k+1} - y_k| < \delta$$

$$\sum_{k=1}^{n} |f(y_k) - f(x_k)| \leq \eta \sum_{k=1}^{n} (y_k - x_k) < \eta.(c - a)$$

By the construction of the intervals and absolute continuity of f,

$$\sum_{k=0}^{n} |f(x_{k+1}) - f(y_k)| < \epsilon$$

Hence,

$$|f(c) - f(a)| = \left| \sum_{k=0}^{n} [f(x_{k+1}) - f(y_k)] + \sum_{k=1}^{n} [f(y_k) - f(x_k)] \right| \leq \epsilon + \eta(b-a)$$

Now, the constants ϵ, η were random. Hence, $f(c) - f(a) = 0$.

□

Theorem 1.71. *A function g is an indefinite integral iff it's an absolutely continuous function.*

Proof. By the lemma 1.62, F is indefinite integral implies it is absolutely continuous. Conversely, if F is absolutely continuous on $[a, b]$, then it is bounded variateand hence sum of 2 monotone functions, say $F = F_1 - F_2$, where F_1, F_2 are monotonously increasing. Hence, by theorem 1.56,

$$\int |F'(x)| \leq F_1(b) + F_2(b) - F_1(a) - F_2(a)$$

And F' is integrable. Declare

$$H(x) = \int_a^x F'(t)dt$$

Here H is absolutely continuous and so is $F - H = f$. By the lemma 1.66, $f'(x) = F'(x) - H'(x) = 0$ a.e and hence by the previous lemma, f is constant. Thus,

$$F(x) = \int_a^x F'(t)dt + F(a)$$

□

Corollary 1.72. *Every absolutely continuous function is an indefinite integral of it's derivative.*

□

- *As an immediate consequence, we have the integration by parts, which is used in proving the Rademacher's theorem in this document.*
  Let f be absolutely continuous function on $[a, b]$ and $\phi \in C_c^1(a, b)$. Then,

  $$-\int_a^b f(x)\phi'(x)dx = \int_a^b f'(x)\phi(x)dx$$

  Proof.

  $$LHS \equiv -\int_a^b f(x)\phi'(x)dx = -\int_a^b \phi'(x)\left[f(a) + \int_a^x f'(t)dt\right]dx$$

$$
\begin{aligned}
&= -f(a)\int_a^b \phi'(x)dx - \int_a^b \phi'(x)\Big(\int_a^x f'(t)dt\Big)dx \\
&= -f(a)\Big(\phi(b)-\phi(a)\Big) - \int_a^b \phi'(x)\Big(\int_a^b \mathcal{X}_{[a,x]}(t)f'(t)dt\Big)dx \\
&= -\int_a^b f'(t)\Big(\int_a^b \phi'(x)\mathcal{X}_{[a,x]}(t)dx\Big)dt \\
&= -\int_a^b f'(t)\Big(\int_a^t \phi'(x)dx\Big)dt \\
&= \int_a^b f'(t)\phi(t)dt \equiv RHS
\end{aligned}
$$

Here, we have used the fact that $\phi(a) = \phi(b) = 0$ and this proves the integration by parts.

□

Definition 1.73. A function $\phi : (a,b) \to \mathbb{R}$ is said to be convex if $\forall x, y \in (a,b), c \in [0,1]$ we have,

$$\phi(cx + (1-c)y) \leq c\phi(x) + (1-c)\phi(y)$$

Lemma 1.74. *If f is convex on (a,b) and if $a \leq x_1 < x_2 < x_3 < x_4 \leq b$, then*

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_4)-f(x_2)}{x_4-x_2} \leq \frac{f(x_4)-f(x_3)}{x_4-x_3}$$

Proof. $x_1 < x_2 < x_3 \implies \exists t \in [0,1]$ such that

$$
\begin{aligned}
x_2 &= tx_3 + (1-t)x_1 \equiv x_1 + t(x_3 - x_1) \\
\implies t &= \frac{x_2-x_1}{x_3-x_1}\ ,\ (1-t) = \frac{x_3-x_2}{x_3-x_1} \\
\implies x_2 &= \frac{x_2-x_1}{x_3-x_1}x_3 + \frac{x_3-x_2}{x_3-x_1}x_1 \\
\implies f(x_2) &\leq \frac{x_2-x_1}{x_3-x_1}f(x_3) + \frac{x_3-x_2}{x_3-x_1}f(x_1) \\
\implies f(x_2) - f(x_1) &\leq \frac{x_2-x_1}{x_3-x_1}\big(f(x_3)-f(x_1)\big) \\
\implies \frac{f(x_2)-f(x_1)}{x_2-x_1} &\leq \frac{f(x_3)-f(x_1)}{x_3-x_1}
\end{aligned}
$$

Also

$$
\begin{aligned}
f(x_2) &\leq \left(1 - \frac{x_3-x_2}{x_3-x_1}\right)f(x_3) + \frac{x_3-x_2}{x_3-x_1}f(x_1) \\
&= f(x_3) + \frac{x_3-x_2}{x_3-x_1}\Big(f(x_1) - f(x_3)\Big)
\end{aligned}
$$

$$\implies \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

Thus

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

Similarly,

$$\frac{f(x_3)-f(x_2)}{x_3-x_2} \leq \frac{f(x_4)-f(x_2)}{x_4-x_2} \leq \frac{f(x_4)-f(x_3)}{x_4-x_3}$$

This proves the lemma.

□

Corollary 1.75. *At all $x \in (a,b)$,*

$$f'_-(x) := \lim_{h\to 0} \frac{f(x)-f(x-h)}{h}$$

$$f'_+(x) := \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

exists and the properties are

- $x \to f'_-(x)$ *is non-decreasing.*
- $x \to f'_+(x)$ *is non-decreasing.*
- $f'_-(x) \leq f'_+(x)$.
- *For almost every $x \in (a,b)$ that is, at the points of continuity of f'_- or f'_+,*

$$f'_-(x) = f'_+(x)$$

Proof. Let $0 < h_1 < h_2$ with $x-h_2 \in (a,b)$, then $x-h_1 \in (a,b)$ and $x-h_2 < x-h_1 < x$. Now, by the convexity of f, we have

$$\frac{f(x)-f(x-h_2)}{h_2} \leq \frac{f(x)-f(x-h_1)}{h_1}$$

$$\implies h \to \frac{f(x)-f(x-h)}{h} \text{ is non increasing, bounded function.}$$

Hence

$$f'_-(x) = \lim_{h\to 0} \frac{f(x)-f(x-h)}{h} \text{ exists.}$$

Let $x < y$ and choose $0 < h_1 < h_2$ such that $x-h_1 \in (a,b)$ and $x < y-h_2$. Then, $x-h_1 < x < y-h_2 < y$ and by the convexity,

$$\frac{f(x)-f(x-h_1)}{h_1} \leq \frac{f(y)-f(y-h_2)}{h_2}$$

Let $h_1, h_2 \to 0$ to get

$$f'_-(x) \leq f'_-(y)$$

This proves the first part of the corollary. The second part of the corollary involves the same proof as the first part.
Let $h_1, h_2 > 0$ such that $x - h_1, x + h_2 \in (a, b)$, then $x - h_1 < x < x + h_2$. By the convexity of f,

$$\frac{f(x) - f(x - h_1)}{h_1} \leq \frac{f(x + h_2) - f(x)}{h_2}$$

Letting $h_1, h_2 \to 0$ gives the third point of the corollary.
Let $x, y \in (a, b)$ with $x < y$ and $h_1 > 0$, $h_2 > 0$ such that

$$x < x + h_1 < y - h_2 < y$$

By the convexity of f,

$$\frac{f(x + h_1) - f(x)}{h_1} \leq \frac{f(y) - f(y - h_2)}{h_2}$$

Let $h_1, h_2 \to 0$ to get

$$f'_+(x) \leq f'_-(y)$$

$$\implies f'_-(x) \leq f'_+(x) \leq f'_-(y) \ , \ \forall x < y$$

Noting that f'_- is monotone and every monotone function is continuous except on a countable set, which has measure zero, let x be the point of continuity of f'_-,

$$\implies \lim_{h \to 0} f'_-(x + h) = f'_-(x)$$

Since $x < x + h$, we have

$$f'_-(x) \leq f'_+(x) \leq f'_-(x + h)$$

Since, $\lim_{h \to 0} f'_-(x + h) = f'_-(x)$, letting $h \to 0$,

$$f'_-(x) \leq f'_+(x) \leq \lim_{h \to 0} f'_-(x + h) = f'_-(x)$$

$$\implies f'_-(x) = f'_+(x) \ ; \text{ for all } x \text{ where } f'_- \text{ is continious}$$

As the complement of the set where f'_- is continuous is of measure zero, this proves the corollary's fourth point.

$\square$

Definition 1.76. Let ϕ be a convex function on (a,b) and let $\widetilde{x} \in (a, b)$.
The line $\{(x, y) \ ; \ y = m(x - \widetilde{x}) + \phi(\widetilde{x})\}$ through $\big(\widetilde{x}, \phi(\widetilde{x})\big)$ is called the supporting line

at $\widetilde{x}$ if it always lies below the graph of ϕ that is, $\phi(x) \geq m(x - \widetilde{x}) + \phi(\widetilde{x})$.

Remark 1.77. Existence of a supporting line.
Let $f : (a, b) \to \mathbb{R}$ be a convex function and $x_0 \in (a, b)$. Then there is a $m\big[\equiv f'_{\pm}(x_0)\big]$ such that

$$f(x) \geq f(x_0) + m.(x - x_0)$$

Proof. Let $L(x) := f(x_0) + f'_{-}(x_0)(x - x_0)$.
Case (1) :

$$a \leq x < x_0$$

Let $h > 0$ be such that $a \leq x < x_0 - h < x_0$. Then, by the convexity of f,

$$\frac{f(x_0) - f(x)}{x_0 - x} \leq \frac{f(x_0) - f(x_0 - h)}{h}$$

Letting $h \to 0$, we have

$$\frac{f(x_0) - f(x)}{x_0 - x} \leq \lim_{h \to 0} \frac{f(x_0) - f(x_0 - h)}{h} \equiv f'_{-}(x_0)$$

$$\Longrightarrow f(x) \geq f(x_0) + f'_{-}(x_0)(x - x_0)$$

Case (2) : Let $x_0 < x \leq b$. Let $h > 0$ be such that $a \leq x_0 - h$. Then $x_0 - h < x_0 < x$ and the convexity of f gives

$$\frac{f(x_0) - f(x_0 - h)}{h} \leq \frac{f(x) - f(x_0)}{x - x_0}$$

Letting $h \to 0$,

$$\Longrightarrow f'_{-}(x_0) \leq \frac{f(x) - f(x_0)}{x - x_0}$$

$$\Longrightarrow f(x) \geq f(x_0) + f'_{-}(x_0)(x - x_0)$$

Combining the above 2 cases, we have

$$f(x) \geq f(x_0) + f'_{-}(x_0)(x - x_0) \quad \forall x \in (a, b)$$

Similarly,

$$f(x) \geq f(x_0) + f'_{+}(x_0)(x - x_0) \quad \forall x \in (a, b)$$

This shows the existence of a supporting line and hence the remark is proved. $\square$

Lemma 1.78. ***Jensen Inequality :***
Let ϕ be a convex function on $\mathbb{R}$. For f an integrable function on $[0,1]$,

$$\int_{[0,1]} \phi(f(t))dt \geq \phi\Big(\int_{[0,1]} f(t)dt\Big)$$

Proof. Let $\alpha = \int f(t)dt$ and $y = m(x-\alpha) + \phi(\alpha)$ be the equation of a supporting line at α. Then, clearly,

$$\phi(f(t)) \geq m(f(t)-\alpha) + \phi(\alpha)$$

Now integrate both the sides to conclude.

□

Lemma 1.79. ***The generalised Jensen's inequality:***
Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function. Let $(X, \mathcal{A}, \mu)$ be a probablitistic measure space, i.e $\mu(X) = 1$. Let $f \in L^1(X, \mathcal{A}, \mu)$, i.e $f : X \to \mathbb{R}$ measurable along with the property $\int_X |f(v)|d\mu(v) < \infty$, then

$$\int_X \phi(f(t))dt \geq \phi\Big(\int_X f(t)dt\Big)$$

Proof. Given that ϕ is convex on $\mathbb{R}$, let $y, m \in \mathbb{R}$ such that $\forall z \in \mathbb{R}$,

$$\phi(z) \geq \phi(y) + m(z-y)$$

Let $z = f(v)$, then

$$\phi(f(v)) \geq \phi(y) + m(f(v)-y)$$

Integrate to get

$$\int_X \phi(f(v))d\mu(v) \geq \phi(y) + m(\int_X f(v)d\mu(v) - y)$$

Choose $y = \int_X f(v)d\mu(v)$, then

$$\int_X \phi(f(v))d\mu(v) \geq \phi(\int_X f(v)d\mu(v))$$

□

Remark 1.80. Let $\phi \in C^2(a,b)$. ϕ is convex iff it's second derivative is globally non negative i.e $\phi'' \geq 0$.

Proof. of the remark: Given that $\phi \in C^2(a,b) \implies \phi'$ exists and by the previous theorem, $\phi'(x)$ is increasing. Hence $\phi''(x) \geq 0$.
Conversely, to show if $\phi'' \geq 0 \implies \phi$ is convex.
Let $a < x_1 < x_2 < b$ and for $t \in [0,1]$, define

$$F(t) = \phi(tx_1 + (1-t)x_2) - t\phi(x_1) - (1-t)\phi(x_2)$$

Hence,

$$F(0) = \phi(x_2) - \phi(x_2) = 0 = \phi(x_1) - \phi(x_1) = F(1)$$

$$F'(t) = \phi'(tx_1 + (1-t)x_2)(x_1 - x_2) - \phi(x_1) + \phi(x_2)$$

$$F''(t) = \phi''(tx_1 + (1-t)x_2)(x_1 - x_2)^2$$

Claim : $F(t) \leq 0\ \forall t \in [0,1]$
Proof of the claim : If the claim is not true, then $\exists t_0 \in (0,1)$ such that

$$F(t_0) = \max_{t \in [0,1]} F(t) > 0$$

Hence,

$$F'(t_0) = 0$$

$$F''(t_0) \leq 0$$

Subclaim : Suppose that $\phi''(x) > 0$, for all x, then $F(t) \leq 0$ for all $t \in (0,1)$.
Proof of the subclaim: If not, by the hypothesis,$\phi''(x) > 0,\ \forall x \in (a,b)$ which implies $F''(t) > 0 \forall t \in [0,1]$. Hence at t_0, $F''(t_0) > 0$ which is a contradiction.
Now, let $\epsilon > 0$. Define

$$\phi_\epsilon(x) = \phi(x) + \frac{\epsilon}{2}x^2$$

Then $\phi''_\epsilon(x) = \epsilon > 0$. Hence, by the subclaim, ϕ_ϵ is convex. i.e

$$\phi_\epsilon\big(tx_1 + (1-t)x_2\big) \leq t\phi_\epsilon(x_1) + (1-t)\phi_\epsilon(x_2)$$

So,

$$\phi\Big(tx_1 + (1-t)x_2\Big) + \frac{\epsilon}{2}\Big(tx_1 + (1-t)x_2\Big)^2$$

$$\leq t\Big[\phi(x_1) + \frac{\epsilon}{2}{x_1}^2\Big] + (1-t)\Big[\phi(x_2) + t{x_2}^2\Big]$$

Tend $\epsilon \to 0$ to get ϕ is convex.

$\square$

1.5.4 Fubini and Tonelli theorem.

Let $(X, \mathbb{A}, \mu)$ and $(Y, \mathbb{B}, v)$ be 2 complete measure spaces. Consider the product of the 2 spaces, denoted by $X \times Y := \{(x, y)\ ;\ x \in X,\ y \in Y\}$.

Definition 1.81.

- $A \times B$ is called a rectangle if $A \subset X$ and $B \subset Y$.
- $A \times B$ is called a measurable rectangle if $A \in \mathbb{A}$ and $B \in \mathbb{B}$.

Remark 1.82. The collection of measurable rectangles $\mathcal{R}$ is a semi algebra as

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

$$(A \times B)^c = (A^c \times B)\ \cup\ (A \times B^c)\ \cup\ (A^c \times B^c)$$

Definition 1.83. For a measurable rectangle $A \times B$, declare the product measure

$$\lambda(A \times B) = \mu(A).v(B)$$

Lemma 1.84. *Let $\{A_i \times B_i\}$ be countable disjoint collection of measurable rectangles whose union is a measurable rectangle $A \times B$. Then*

$$\lambda(A \times B) = \sum \lambda(A_i \times B_i)$$

Proof. $\{A_i \times B_i\}$ are disjoint and hence,

$$\mathcal{X}_A(x)\mathcal{X}_B(x) = \mathcal{X}_{A\times B}(x, y) = \sum_{i=1}^{\infty} \mathcal{X}_{A_i \times B_i}(x, y) = \sum_{i=1}^{\infty} \mathcal{X}_{A_i}(x)\mathcal{X}_{B_i}(y)$$

Fix $x \in A$.

$$y \to \mathcal{X}_{A_i}(x)\mathcal{X}_{B_i}(y) \text{ is measurable}$$

Then for each $y \in B$, $(x, y) \in A_i \times B_i$ for exactly one $A_i \times B_i$ This tells that B is a disjoint union of those B_i such that the corresponding A_i has x. Thus, as v is countable additive,

$$\sum vB_i.\mathcal{X}_{A_i}(x) = vB.\mathcal{X}_A(x)$$

By MCT,

$$\sum \int vB_i.\mathcal{X}_{A_i}(x)d\mu = \int vB.\mathcal{X}_A(x)d\mu$$

Hence,

$$\sum vB_i.\mu A_i = vB.\mu A$$

$\square$

Remark 1.85. This lemma implies that there is a unique extension of λ to a measure on the algebra $\mathcal{R}'$ which contains all finite disjoint union of sets in $\mathcal{R}$. Now extend λ to a complete measure on a sigma algebra $\mathcal{S}$ containing $\mathcal{R}$ and denote it by $\mu \times v$.

Definition 1.86. For any $E \subset X \times Y$, $x \in X$, $y \in Y$,

- Define $E_x = \{y \in Y : (x, y) \in E\}$.
- Define $E^y = \{x \in X : (x, y) \in E\}$.

Remark 1.87. $\mathcal{X}_{E_x}(y) = \mathcal{X}_E(x, y)$ and $\mathcal{X}_{E^y}(x) = \mathcal{X}_E(x, y)$.

Remark 1.88. Note that proving a statement for E_x is similar to proving the analogous statement for E^y and hence, for the following lemmas and theorems in this subsection, only one case (say E_x) will be proved.

Lemma 1.89. *Let $x \in X$ and $E \subset \mathcal{R}_{\sigma\delta}$. Then E_x is a measurable subset of Y and E^y is a measurable subset of X.*

Proof. Case 1 : If $E \in \mathcal{R}$, then by the definition, E_x is measurable.
Case 2 : If $E \in \mathcal{R}_\sigma$, then $E = \cup_{i=1}^{\infty} E_i$ such that each E_i is a measurable rectangle. Now

$$E_x = \left(\cup_{i=1}^{\infty} E_i\right)_x = \cup_{i=1}^{\infty} (E_i)_x$$

Now $(E_i)_x$ is measurable for all i and hence, E_x is measurable.
Case 3: If $E \in \mathcal{R}_{\sigma\delta}$, then $E = \cap_{i=1}^{\infty} E_i$ such that each $E_i \in \mathcal{R}_\sigma$. Now

$$E_x = \left(\cap_{i=1}^{\infty} E_i\right)_x = \cap_{i=1}^{\infty} (E_i)_x$$

Now $(E_i)_x$ is measurable for all i, by the case 2, and hence, E_x is measurable.

$\square$

Lemma 1.90. *For $E \in \mathcal{R}_{\sigma\delta}$ with $\mu \times v(E) < \infty$, g defined by*

$$g(x) = vE_x$$

is a measurable function of x and

$$\int g d\mu = \mu \times v(E)$$

Proof. Case 1 : If $E = A \times B$ is a measurable rectangle, then $g(x) = v(B)$ if $x \in A$, else, it is an empty set and hence measurable. Also $\int g d\mu = \mu(A)v(B) = \mu \times v(E)$.
Case 2 : If E is in $\mathcal{R}_\sigma$. W.L.O.G, assume that $E = \cup E_i$ where E_i's are disjoint. Declare

$$g_i(x) = v\big[(E_i)_x\big]$$

g is thus measurable and is a positive function. Also note that $g = \sum g_i$ and hence g is measurable. By M.C.T,

$$\int g d\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int g_i d\mu = \sum \int g_i d\mu = \sum \mu \times v(E_i) = \mu \times v(E)$$

Case 3 : If E is in $\mathcal{R}_{\sigma\delta}$. W.L.O.G, assume that $E = \cap E_i$ where $E_i \subset E_{i+1}$ and $E_i \in \mathcal{R}_\sigma$. By approximation, we can assume $\mu \times v(E_1) < \infty$. Let $g_i(x) = v[(E_i)_x]$.
Hence, $g = \lim g_i$ and hence g is measurable. Since $\int g_1 d\mu = \mu \times v(E) < \infty$, $g_1(x) < \infty$ can be concluded for a.e x . Now $\{(E_i)_x\}$ is a decreasing sequence of finite measurable sets and their intersection gives E_x. From the knowledge of measure theory, since the first set in the intersection has finite measure, we can have

$$g(x) = v(E_x) = \lim v\big[(E_i)_x\big] = \lim g_i(x)$$

And thus

$$g_i \to g$$

a.e. which also implies g is measurable.
Since $g_1 \geq g_i \geq 0$, the D.C.T tells

$$\int g d\mu = \lim \int g_i d\mu = \lim \mu \times v(E_i) = \mu \times v(E)$$

which is true form the limit of the intersection property.

$\square$

Lemma 1.91. *For E with $\mu \times v(E) = 0$, we have for μ almost every x , $v(E_x) = 0$.*
Remark : *Note that the product measure is defined only for rectangles until now. To make sense of $\mu \times v(E)$ with $E \subset X \times Y$, define $\mathcal{C}$ to be the algebra generated by the semi algebra $\mathcal{R}$ and hence define,*

$$\lambda^*(E) := \inf\left\{\textstyle\sum_{i=1}^{\infty} \mu \times v(A_i) \ ; \ E \subset \bigcup_{i=1}^{\infty} A_i \ , \ \{A_i\}_{i=1}^{\infty} \subset \mathcal{C}\right\}$$

By the Caratheodory's theorem, there is a complete sigma algebra containing $\mathcal{R}$(as well as $\mathcal{C}$) and a measure λ on the sigma algebra, such that λ matches with the $\mu \times v$ on $\mathcal{R}$. Hence, from now on, if the set is not mentioned to be in $\mathcal{R}$ and is a subset of $X \times Y$, then $\mu \times v$ is to be taken as λ^*.

Proof. Using the approximations dictated by the infimum property of λ^*, by the unions and intersections and controlling measure with it, we can get a $G \in \mathcal{R}_{\sigma\delta}$ such that $E \subset G$ and $\mu \times v(G) = 0$. And by the previous lemma, we get,

$$0 = \mu \times v(G) = \int v(G_x) d\mu(x)$$

As $x \to v(G_x)$ is a positive function, the above implies that for μ almost every x, $v(G_x) = 0$. Also $E_x \subset G_x$ and so $v(E_x) = 0$ for $\mu-$ a.e x.

□

Lemma 1.92. *For E to be a measurable subset of $X \times Y$ with $\mu \times v(E) < \infty$, for almost every x, E_x is a measurable subset of Y. Also,*

$$g(x) = v(E_x)$$

is measurable function for μ almost every x and

$$\int g(x) d\mu(x) = \mu \times v(E)$$

Proof. By the approximations by the unions and intersections and controlling measure with it, $\exists G \in \mathcal{R}_{\sigma\delta},\ E \subset G$ with $\mu \times v(E) = \mu \times v(G)$. Now, look at $F = G - E$. Since E and G are measurable, so is F.

$$\mu \times v(G) = \mu \times v(E) + \mu \times v(F)$$

Since $\mu \times v(E)$ is finite and is equal to $\mu \times v(G)$, we can conclude

$$\mu \times v(F) = 0$$

Thus by the previous lemma, we can conclude, $v(F_x) = 0$ for μ almost all x. Hence

$$E_x \subset G_x = F_x \cup E_x \implies v(E_x) \le v(G_x) \le v(F_x) + v(E_x) = v(E_x) \text{ for } \mu \text{ a.e } x$$

$$\implies g(x) := v(E_x) = v(G_x) \text{ for } \mu \text{ a.e } x$$

This tells that g is a measurable function, by the lemma 1.90. And also, by the same lemma (1.90), we have

$$\int g d\mu = \mu \times v(G) = \mu \times v(E)$$

□

Theorem 1.93. ***(Fubini theorem) :***
For $(X, \mathbb{A}, \mu)$ and $(Y, \mathbb{B}, v)$ be 2 complete measure spaces as before and f is an integrable function on $X \times Y$, we have

- *For almost every x, the function f_x defined by $f_x(y) = f(x, y)$ is an integrable function on Y.*
- *For almost every y, the function f_x defined by $f_y(x) = f(x, y)$ is an integrable function on X.*
- *$\int_Y f(x, y) dv(y)$ is an integrable function on X.*
- *$\int_X f(x, y) d\mu(x)$ is an integrable function on Y.*
- $$\int_X \left[\int_Y f dv \right] d\mu = \int_{X \times Y} f d(\mu \times v) = \int_Y \left[\int_X f d\mu \right] dv$$

Proof. It suffices to prove the theorem for one of the cases, because of the symmetry involved between x and y. Also note that for any f, it can be decomposed as $f^+ - f^-$ and if the conclusion of the theorem holds for 2 functions implies it holds even for their difference. Hence, W.L.O.G assume f is positive. Now f can be approximated by increasing simple functions $\{\phi_n\}$ such that it take finite values in the image and vanished outside a set of finite meaasure. If the theorem is true for charecteristic function over a set of finite measure, by M.C.T we can conclude for the positive f. The previous lemma tells that the theorem is true for a charecteristic function over a set of finite measure. Hence by M.C.T, we have

$$\int_Y f(x, y) dv(y) = \lim_{n \to \infty} \int_Y \phi_n(x, y) dv(y)$$

Also, by M.C.T,

$$\int_X \left[\int_Y f dv \right] d\mu = \lim_{n \to \infty} \int_X \left[\int_Y \phi_n dv \right] d\mu = \lim_{n \to \infty} \int_{X \times Y} \phi_n d(\mu \times v) = \int_{X \times Y} f d(\mu \times v)$$

□

Remark 1.94. Here, the sigma finiteness of the spaces are not required.

Theorem 1.95. ***(Tonelli theorem) :***
For $(X, \mathbb{A}, \mu)$ and $(Y, \mathbb{B}, v)$ be two $\sigma-$finite measure spaces and f is a non-negative measurable function on $X \times Y$, we have

- *For almost every x, the function f_x defined by $f_x(y) = f(x, y)$ is an non-negative measurable function on Y.*
- *For almost every y, the function f_x defined by $f_y(x) = f(x, y)$ is an non-negative measurable function on X.*
- *$\int_Y f(x, y)dv(y)$ is an measurable function on X.*
- *$\int_X f(x, y)d\mu(x)$ is an measurable function on Y.*
- $$\int_X \left[\int_Y f dv \right] d\mu = \int_{X \times Y} f d(\mu \times v) = \int_Y \left[\int_X f d\mu \right] dv$$

Proof. As the spaces are $\sigma-$finite , any non negative function can be approximated by simple integrable functions and hence the similar proof like that in the fubini theorem works.

□

1.5.5 Change of Variables

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map and $\{e_1, e_2 \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$.
Then T is represented by a matrix with the elements as $T_{i,j} = < Te_j, e_i >$. By the property of determinant, determinant is the product of the eigenvalues. Let $GL_n(\mathbb{R})$ to be the group of all $n \times n$ invertible matrices.

Theorem 1.96. *Let $T \in GL_n(\mathbb{R})$, then*

- *If f is measurable and either $f \geq 0$ or $f \in L^1(\mathbb{R}^n)$, then so is $f \circ T$ and*
$$\int_{\mathbb{R}^n} f(x)dx = |detT| \int_{\mathbb{R}^n} (f \circ T)(x)dx$$
- *If E be a measurable subset, then $T(E)$ is lebesgue measurable and*
$$\mathcal{L}^n(T(E)) = |detT|\mathcal{L}^n(E)$$

Proof. Note : Any $T \in GL_n(\mathbb{R})$ can be written as product of finitely many transformations of elementary matrices of the type:

- I : $T_1(x_1, x_2, \ldots, x_i, \ldots x_n) = (x_1, x_2, \ldots, cx_i, \ldots x_n)$; $c \neq 0$
- II : $T_2(x_1, x_2, \ldots, x_i, \ldots x_n) = (x_1, x_2, \ldots, x_i + cx_k, \ldots x_n)$, $k \neq i$
- III : $T_3(x_1, x_2, \ldots, x_i, \ldots, x_j \ldots x_n) = (x_1, x_2, \ldots, x_j, \ldots, x_i \ldots x_n)$

Now, Since T is continuous , f is borel measurable implies $f \circ T$ is Borel measurable. For the case (I), with the usage of fubini's theorem,

$$\int_{\mathbb{R}^n} f \circ T(x)dx = \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, x_2, \ldots, cx_i, \ldots, x_n)dx_1 dx_2 \ldots dx_{i-1} dx_i dx_{i+1} \ldots dx_n$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2, \ldots, cx_i, \ldots, x_n)dx_i \right) dx_1 dx_2 \ldots dx_{i-1} dx_{i+1} \ldots dx_n$$

$$= \frac{1}{|c|} \int_{\mathbb{R}^n} f(x)dx$$

And

$$|c| \int_{\mathbb{R}^n} f \circ T(x)dx = \int_{\mathbb{R}^n} f(x)dx$$

Where $|detT| = |c|$. For the case (II), W.L.O.G assume, that the change is happening at the nth place.

$$\int_{\mathbb{R}^n} f \circ T(x)dx = \int_{\mathbb{R}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, x_2, \ldots, x_n + cx_i)dx_1 dx_2 \ldots dx_n$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \ldots \left[\int_{\mathbb{R}} f(x_1, x_2, \ldots, x_n + cx_i)dx_n \right] dx_{n-1} \ldots dx_1$$

x_i is constant for the integration over x_n and $\int g(a+t)dt = \int g(t)dt$ due to translation invariant lebesgue measure. Hence

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \ldots \left[\int_{\mathbb{R}} f(x_1, x_2, \ldots x_n)dx_n \right] dx_{n-1} \ldots dx_1 = \int_{\mathbb{R}^n} f(x)$$

Also note that $|detT_2| = 1$. This implies the justification of the first point of the hypothesis for the case (II). For the case (III): Similar calculations as the previous cases, but with $detT_3 = -1$ and hence $|detT_3| = 1$.

For the general case of $T \in GL_n(\mathbb{R})$,

$$T = T_1 \circ \cdots \circ T_k$$

where T_i's are like in one of the cases. So

$$\int f(x)dx = |detT_1| \int f \circ T_1(x)dx$$

$$= |detT_1||detT_2| \int f \circ T_1 \circ T_2(x)dx$$

$$\vdots$$

$$= |detT| \int f \circ T dx$$

To prove the second point of the theorem, consider $f = \mathcal{X}_{T(E)}$.

$\square$

Theorem 1.97. ***Change of variables formula:*** *Suppose Ω is an open subset of $\mathbb{R}^n$ and $g : \Omega \to g(\Omega)$, a C^1- diffeomorphism.*

- *If f is lebesgue measurable on $g(\Omega)$, then $f \circ g$ is lebesgue measurable on Ω and if $f \geq 0$ or $f \in L^1$ then*

$$\int_{g(\Omega)} f(x)dx = \int_{\Omega} (f \circ g)(x)|detD_x g|dx$$

- *If $E \subset \Omega$ is lebesgue measurable, so is g(E) and*

$$m(g(E)) = \int_E |detD_x g|dx$$

where $m(A)$ denotes the lebesgue measure of $A \subset \mathbb{R}^n$.

Proof. First we prove the result for the borel measurable functions and borel sets. Since g and g^{-1} are both continuous , $f \circ g$ is borel measurable and $g(E)$ is a borel set.
Let Q be a cube with center x and side $2h$ such that $Q \subset \Omega$, $Q = \{y \; ; \; ||x - y||_\infty \leq h\}$.
Here

$$||x - y||_\infty = \max_{1 \leq j \leq n} |x_j - y_j|$$

Now, $g_j(y) - g_j(x) = Dg_j(z).(y - x)$ where z is in the line joining x and y.
So for any $y \in Q$,

$$||g_j(y) - g_j(x)||_\infty \leq h \bigg(\sup_{z \in Q} ||Dg(z)|| \bigg)$$

i.e $g(Q)$ is contained in a cube of side length $\big(\sup_{z \in Q} ||Dg(z)||\big)$ times that of Q.
Therefore

$$m\bigg(g(Q)\bigg) \leq \bigg(\sup_{z \in Q} ||Dg(z)||\bigg)^n m(Q)$$

Let $T \in GL_n(\mathbb{R}^n)$. Now apply the previous change of variables formula to get

$$m\Big(g(Q)\Big) = m\Big(T \circ T^{-1} \circ g(Q)\Big)$$

$$= |detT| m\Big(T^{-1} \circ g(Q)\Big)$$

$$\leq |detT| \sup_{z \in Q} \Big[||D(T^{-1} \circ g)(z)||_\infty\Big]^n m(Q)$$

Since $g^{-1} \circ g(x) = x$ and $D(g^{-1} \circ g) = I$ i.e $Dg^{-1}(g(x)).Dg(x) = I$

Fix $\epsilon > 0$. Subdividing Q into subcubes $Q_1, Q_2, \ldots Q_r$ whose centres are $x_1, x_2, \ldots, x_r$ and their interiors are disjoint with side length less than δ, where δ corresponds to

$$\begin{bmatrix} ||Dg^{-1}\big(g(x_i)\big)Dg(y)||^n < 1 + \epsilon \\ \text{whenever } ||y - x||_\infty < \delta \end{bmatrix} \text{ ; by the continuity of } Dg \text{ and } Dg^{-1}$$

By the previous estimate on each Q_i with $T = Dg(x_i)$, we get

$$m(g(Q_i)) \leq |det(Dg(x_i))| \left[\sup_{y \in Q_i} ||Dg^{-1}(x_i).Dg(y)||_\infty\right]^n m(Q_i)$$

Then,

$$m(g(Q)) \leq \sum_{i=1}^{r} m\Big(g(Q_i)\Big)$$

$$= (1 + \epsilon) \sum_{i=1}^{r} \Big|det\big(Dg(x_i)\big)\Big| m(Q_i)$$

Note that $\sum_{i=1}^{r} |detDg(x_i)|.m(Q_i)$ is Riemann sum corresponding to the integral $\int_Q |detDg|$ corresponding to the partition $Q = \bigcup_{i=1}^{r} Q_i$.

Therefore as $\delta \to 0$, we get

$$m\Big(g(Q)\Big) \leq (1 + \epsilon) \int_Q |detDg|$$

Now $\epsilon > 0$ was arbitrary. Hence

$$m\Big(g(Q)\Big) \leq \int_Q |detDg| dx$$

The above, now holds for any open cube contained in Ω. Noting that any open set is a disjoint union of countable cubes. And hence,

$$m\Big(g(U)\Big) \leq \int_U |detDg| dx \text{ ; } \forall U \subset \Omega \text{ , } U \text{ is open.}$$

Let E be a borel set of finite measure, there is a sequence of open sets by the outer regularity such that

- $U_1 \supset U_2 \supset U_3 \supset \dots$
- $E \subset \bigcap_{i=1}^{\infty} U_i$
- $m\big(\cap_{i=1}^{\infty} (U_i - E)\big) = 0$

Now,

$$m\bigg(g(E)\bigg) \le m\bigg(g(\cap_{i=1}^{\infty} U_i)\bigg) = \lim_{n\to\infty} m\bigg(g(U_n)\bigg)$$

$$= \lim_{n\to\infty} \int_{U_n} |detDg| = \int_{\cap U_n} |detDg|dx$$

$$= \int_E |detDg|dx$$

Now if E is borel , then E can be decomposed as

- $E = \bigcup_{n=1}^{\infty} E_n$
- $m(E_n) < \infty$
- $E_1 \subset E_2 \subset E_3 \subset \dots$

Then,

$$m\bigg(g(E)\bigg) \le \int_E |detDg|dx \text{ ; } \forall E \subset \Omega,\ E \text{ is Borel.}$$

Consider the function,

$$f = \sum_{i=1}^{n} \alpha_i \mathcal{X}_{A_i} \quad ; \quad \alpha_i \ge 0$$

to be simple function on $g(\Omega)$, then

$$\int_{g(\Omega)} f(x)dx = \sum_{i=1}^{n} \alpha_i m(A_i)$$

$$\le \sum_{i=1}^{n} \alpha_i \int_{g^{-1}(A_i)} |detDg| \text{ ; as } A_i = g \circ g^{-1}(A_i)$$

$$= \int_{\Omega} \sum \alpha_i \mathcal{X}_{g^{-1}(A_i)} |detDg|$$

$$= \int_{\Omega} (f \circ g)(x) |detDg|$$

By MCT, for any $f \ge 0$ and borel measurable,

$$\int_{g(\Omega)} f(x)dx \le \int_{\Omega} (f \circ g)(x)|detDg|dx \text{ ; for all } f \text{ borel}$$

Applying this result to f replaced by $f \circ g$ and for g to be g^{-1}, we have,

$$\int_{\Omega} (f \circ g)(x)|detDg|dx \leq \int_{g(\Omega)} (f \circ g \circ g^{-1})\ |detD(g \circ g^{-1})(x)|\ |detg^{-1}(x)|$$

$$= \int_{g(\Omega)} f(x)dx$$

This proves the result for $f \geq 0$.
For the case, $f \in L^1$, decompose f as $f^+ - f^-$ and do the same for f^+ and f^-, individually to get the desired result.
For the part (b), consider $f \equiv \mathcal{X}_E$.

□

Chapter 2

Geometric measure theory.

2.1 Covering theorems

In this chapter, a ball generally refers to a closed ball until mentioned otherwise. Also,in this chapter, until otherwise mentioned, measure of a set is same as the outer measure of that set.

Definition 2.1.
- If B is a closed ball in $\mathbb{R}^n$, say around x, denote $\hat{B}$ to be that ball around x but with radius 5 times that of B
- A collection $\mathcal{F}$ of closed balls in $\mathbb{R}^n$ is said to be a cover for a set $A \subset \mathbb{R}^n$ if

$$A \subset \bigcup_{B \in \mathcal{F}} B$$

- $\mathcal{F}$ is said to be a fine cover for A if in addition from being a cover for A, the following holds :

$$\inf\{diamB \ ; \ x \in B, B \in \mathcal{F}\} = 0 \text{ ; for each } x \in A$$

2.1.1 Vitali's Covering lemma .

Theorem 2.2. *Vitali's covering theorem :*
Let $\mathcal{F}$ be any collection of non-degenerate (radius to be non zero) closed balls in $\mathbb{R}^n$ along with the criteria that

$$\sup\{diamB \ ; \ B \in \mathcal{F}\} < \infty$$

Then there exists a countable family $\mathcal{G}$ of disjoint closed balls in $\mathcal{F}$ such that

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

Proof. Declare

$$D \equiv \sup\{diamB \ ; \ B \in \mathcal{F}\}$$

For the D defined as above, which is finite by the hypothesis, declare

$$\mathcal{F}_j \equiv \{B \in \mathcal{F} \ ; \ \frac{D}{2^j} < diamB \leq \frac{D}{2^{j-1}}\} \ ; \text{ for all } j = 1, 2 \dots$$

Algorithm to construct $\mathcal{G}_i$:
(step 1) : Choose $\mathcal{G}_1$ to be the maximal disjoint set of closed balls from $\mathcal{F}_1$.
(Step 2) : Choose $\mathcal{G}_2$ to be the maximal collection of disjoint sets from $\mathcal{F}_2$ such that

$$\forall \ B \in \mathcal{G}_1 \ , \ \forall \ B' \in \mathcal{G}_2 \ , \ B \cap B' = \phi$$

(Step 3) : For a general $N \in \mathbb{N}$, assume that $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{N-1}$ has been chosen.
Choose $\mathcal{G}_N$ to be the maximal disjoint subcollection of

$$\{B \in \mathcal{F}_N \ ; \ B \cap B' = \phi \ , \text{ for all } B' \in \bigcup_{j=1}^{N-1} \mathcal{G}_j\}$$

Once, $\mathcal{G}_i$'s are chosen as above, declare

$$\mathcal{G} = \bigcup_{i=1}^{\infty} \mathcal{G}_i$$

Clearly $\mathcal{G}$ is a collection of disjoint balls and

$$\mathcal{G} \in \mathcal{F}$$

Claim : For each $B \in \mathcal{F}$, there is a ball $B' \in \mathcal{G}$ such that

- $B \cap B' \neq \phi$
- $B \subset \hat{B}'$

Proof of the claim : Fix $B \in \mathcal{F}$. If $B \in \mathcal{G}_i$ for some i, the claim is trivially true.
Suppose that $B \notin \mathcal{G}_i$, for any i. By the definition of $\{\mathcal{F}_j\}_{j=1}^{\infty}$, there is a number j such that $B \in \mathcal{F}_j$. By the construction of $\mathcal{G}_j$, that is, the maximality of $\mathcal{G}_j$, $B \notin \mathcal{G}_j$ and thus,

there is a ball $B' \in \mathcal{G}_i$ with $B \cap B' \neq \phi$, for some $i \leq j-1$ or it violates the maximal disjointness of $\mathcal{G}_j$, that is, there is some $B' \in \mathcal{G}_j$ such that $B \cap B' \neq \phi$.

$$\implies \exists B' \in \mathcal{G}_i \text{ with } B \cap B' \neq \phi, \text{ for some } i \leq j$$

Now

$$diam(B) \leq \frac{D}{2^{j-1}} \leq \frac{2D}{2^{i-1}} < 2\ diam(B')$$

So, $diamB \leq 2\ diamB' \implies B \subset \hat{B}'$. The claim and hence the theorem is true.

□

Corollary 2.3. *Assume that $\mathcal{F}$ is a fine cover of A by closed balls and*

$$\sup\{diamB\ ;\ B \in \mathcal{F}\} < \infty$$

Then there is a countable family of disjoint balls $\mathcal{G}$ in $\mathcal{F}$ such that for each finite subset $\{B_1, B_2, \ldots, B_n\} \subset \mathcal{F}$,

$$\{A - \bigcup_{k=1}^{n} B_k\} \subset \{\bigcup_{B \in \mathcal{G} - \{B_1, B_2, \ldots, B_n\}} \hat{B}\}$$

Proof. Choose $\mathcal{G}$ as in the proof of the previous theorem.
Fix a finite collection $\{B_1, B_2, \ldots, B_n\} \subset \mathcal{G}$ as in the hypothesis.
Case 1 : If $A \subset \bigcup_{k=1}^{n} B_k$, then there is nothing to show.
Case 2 : Else, let $y \in A - \bigcup_{k=1}^{n} B_k$. Note that the balls B_i's are closed and hence the point y is at a finite, non-zero distance from the boundary of all balls from the set $\{B_1, B_2 \ldots, B_n\}$. $\mathcal{F}$ is given to be a fine cover for A. Hence, for that finite, non zero distance, say d, there is an element B in $\mathcal{F}$ such that

- $diamB < d$
- $B \cap B_j = \phi$; for all $j = 1, 2 \ldots, n$

By the proof of the claim in the proof of the previous theorem, there is $B' \in \mathcal{G}$ such that $B \cap B' \neq \phi$, which tells that $B' \notin \{B_1, B_2, \ldots, B_n\}$. Also, note by the proof of the same claim that $y \in B$, $B \subset \hat{B}'$. This proves the corollary.

□

Corollary 2.4. *Let $U \subset \mathbb{R}^n$ be open . Fix $\delta > 0$. Then there is a countable collection $\mathcal{G}$ of disjoint closed balls in U such that $diam B \leq \delta$, for all $B \in \mathcal{G}$ and*

$$\mathcal{L}^n \left(U - \bigcup_{B \in \mathcal{G}} B \right) = 0$$

Proof. Choose and fix θ such that $1 - \frac{1}{5^n} < \theta < 1$.
Case 1 : Assume $\mathcal{L}^n(U) < \infty$.
Claim : There is a finite collection $\{B_i\}_{i=1}^{M_1}$ of disjoint closed balls in U such that $diam(B_i) < \delta$ for $i = 1, 2 \ldots, M_1$, along with the criteria that

$$\mathcal{L}^n \left(U - \bigcup_{i=1}^{M_1} B_i \right) \leq \theta \mathcal{L}^n(U)$$

Proof of the claim : Declare

$$\mathcal{F}_1 := \{B \text{ is a closed ball} \ ; \ B \subset U \ , \ diam B < \delta\}$$

$\mathcal{F}_1$ is clearly a fine cover and hence, by the vitali's covering theorem, there is a countable disjoint family $\mathcal{G}_1 \subset \mathcal{F}_1$ such that

$$U \subset \bigcup_{B \in \mathcal{G}_1} \hat{B}$$

Thus

$$\mathcal{L}^n(U) \leq \sum_{B \in \mathcal{G}_1} \mathcal{L}^n(\hat{B}) = 5^n \sum_{B \in \mathcal{G}_1} \mathcal{L}^n(B)$$

The above is true by the translation invariant property of the lebesgue measure. Now, because of the disjointness of the elements of $\mathcal{G}_1$,

$$= 5^n \mathcal{L}^n \left(\bigcup_{B \in \mathcal{G}_1} B \right)$$

And thus,

$$\mathcal{L}^n \left(\bigcup_{B \in \mathcal{G}_1} B \right) \geq \frac{1}{5^n} \mathcal{L}^n(U)$$

Subtracting $\mathcal{L}^n(U)$ from both the sides,

$$\mathcal{L}^n \left(U - \bigcup_{B \in \mathcal{G}_1} B \right) \leq \left(1 - \frac{1}{5^n} \right) \mathcal{L}^n(U)$$

Since $\mathcal{G}_1$ is countable, there are balls $B_1, B_2, \ldots, B_{M_1}$ such that

$$\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta \mathcal{L}^n(U)$$

Now, declare

$$U_2 = U - \bigcup_{i=1}^{M_1} B_i$$

$$\mathcal{F}_2 = \{B \text{ is a closed ball } \; ; \; B \subset U_2, \; diam B < \delta\}$$

$\mathcal{F}_2$ is a fine cover and thus applying the same logic as above , we get $B_{M_1+1}, B_{M_1+2}, \ldots, B_{M_2}$ such that

$$\mathcal{L}^n\left(U_2 - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta \mathcal{L}^n(U_2)$$

Hence,

$$\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_2} B_i\right) = \mathcal{L}^n\left(U_2 - \bigcup_{i=M_1+1}^{M_2} B_i\right) \leq \theta \mathcal{L}^n(U_2) \leq \theta^2 \mathcal{L}^n(U)$$

Inductively we get disjoint balls such that

$$\mathcal{L}^n\left(U - \bigcup_{i=1}^{M_k} B_i\right) \leq \theta^k \mathcal{L}^n(U) \; ; \; (k = 1, 2, \ldots)$$

Noting that $\theta < 1$ implies $\theta^k \to 0$, the corollary is true for finite measure sets.
Case 2 : If $\mathcal{L}^n(U) = \infty$, then the trick is to look at

$$U_l = \{x \in U \; ; \; l < |x| < l+1\} \; ; \; (l = 0, 1, 2, \ldots)$$

Note that U_l's are open and by the previous step, denote B_{li} to be the balls generated as per the corollary within U_l. Thus, by the disjointness of sets and noting that the measure of the boundary of U_l, is 0,

$$\mathcal{L}^n\left(U - \bigcup_{l,i=1}^{\infty} B_{li}\right) = \mathcal{L}^n\Big(\bigcup_{l=1}^{\infty} (U_l - \bigcup_{i=1}^{\infty} B_{li})\Big) + \mathcal{L}^n\Big(\bigcup_{l=1}^{\infty} (\partial U_l - \bigcup_{i=1}^{\infty} B_{li})\Big)$$

$$= \mathcal{L}^n\Big(\bigcup_{l=1}^{\infty} (U_l - \bigcup_{i=1}^{\infty} B_{li})\Big) = \sum_{l=1}^{\infty} \mathcal{L}^n\left(U_l - \bigcup_{i=1}^{\infty} B_{li}\right) = 0$$

This proves the corollary.

$\square$

2.1.2 Besicovitch covering lemma .

Theorem 2.5. *If for any collection, say $\mathcal{F}$, consisting of non-degenerate closed balls in $\mathbb{R}^n$ with the criteria*

$$\sup\{\ diamB\ ;\ B \in \mathcal{F}\} < \infty$$

and if A is the centers of the balls from $\mathcal{F}$, then there is a dimensional constant N such that there are families of disjoint balls from $\mathcal{F}$, say $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_N$,(note that the balls within each $\mathcal{G}_i$ are disjoint, but may not be disjoint from a ball from some other family $\mathcal{G}_j$) such that

$$A \subset \bigcup_{i=1}^{N} \bigcup_{B \in \mathcal{G}_i} B$$

Proof. As before, assume for now that A is bounded. Declare

$$D := \sup\{diamB\ ;\ B \in \mathcal{F}\}$$

By the sup property, choose $B_1 := B(a_1, r_1) \in \mathcal{F}$ such that $\frac{3}{4}\frac{D}{2} \leq r_1$.
Given that B_1 is chosen, choose B_2 as :
Case 1 : If $A_2 := A - B_1$ is empty, then stop and set $J = 1$.
Case 2 : If $A_2 := A - B_1$ is not empty , then by the sup property choose $B_2 = B(a_2, r_2) \in \mathcal{F}$ with $a_2 \in A_2$ such that

$$\frac{3}{4}\sup\{r\ ;\ B(a,r) \in \mathcal{F}\ ,\ a \in A_2\} \leq r_2 < \sup\{r\ ;\ B(a,r) \in \mathcal{F}\ ,\ a \in A_2\}$$

Inductively with $B_1, B_2, \ldots, B_{j-1}$ chosen earlier, choose B_j as
Case 1 : If $A_j := A - \bigcup_{i=1}^{j-1} B_i$ is empty, then stop and set $J = j - 1$.
Case 2 : If $A_j := A - \bigcup_{i=1}^{j-1} B_i$ is not empty , then by the sup property choose $B_j = B(a_j, r_j) \in \mathcal{F}$ with $a_j \in A_j$ such that

$$\frac{3}{4}\sup\{r\ ;\ B(a,r) \in \mathcal{F}\ ,\ a \in A_j\} \leq r_j < \sup\{r\ ;\ B(a,r) \in \mathcal{F}\ ,\ a \in A_j\}$$

With the above logic, if $A_k \neq \phi$ for all k , then declare $J = \infty$.

Claim 1 : Suppose $i > j$, then $r_i \leq \frac{4}{3} r_j$.
Proof of the claim 1 : Suppose $i > j$, then observe that $a_i \in A_j$ and $a_j \notin A_i$. Thus

$$\frac{3}{4} r_i \leq \frac{3}{4}\sup\{r\ ;\ B(a,r) \in \mathcal{F}, a \in A_j\} \leq r_j$$

This proves the claim 1.

Claim 2 : The balls $\{B(a_j, \frac{r_j}{3})\}_{j=1}^{J}$ are disjoint.
Proof of the claim 2 : Let $i > j$. Then, as observed before, $a_i \in A_j$ and $a_j \notin A_i$. Hence $a_i \notin B_j$. So

$$|a_j - a_i| > r_j = \frac{r_j}{3} + \frac{2r_j}{3} \geq \frac{r_j}{3} + \left(\frac{2}{3}\right)\left(\frac{3}{4}\right) r_i > \frac{r_j}{3} + \frac{r_i}{3}$$

This tells that the radius of any 2 balls in the set considered , is strictly greater than sum of the individual radii and hence, the balls are disjoint. This proves the claim 2.

Claim 3 : If $J = \infty$, then $\lim_{j \to \infty} r_j = 0$.
Proof of the claim 3 : By the claim 2, the balls $\{B(a_j, \frac{r_j}{3})\}_{j=1}^{J}$ are disjoint. Now, since a_j is in A which is bounded as per the assumption,that is, $A \subset B(0, M)$ for some $M > 0$. Hence, by summing the measures of disjoint balls, $\sum_j \alpha(n) \left(\frac{r_j}{3}\right)^n < M$, where $\alpha(n)$ denotes the lebesgue measure of the unit ball in $\mathbb{R}^n$. Since the tail of the convergent series goes to 0, we can conclude that $r_j \to 0$. This proves the claim 3.

Claim 4 : $A \subset \bigcup_{j=1}^{J} B_j$.
Proof of the claim 4 : If $J < \infty$, then, by the construction of B_i, the inclusion is trivial. Hence, assume $J = \infty$.
Let $a \in A$. Then, by the definition of A, there is some $r > 0$ such that $B(a, r) \in \mathcal{F}$. If $a \notin \bigcup_{j=1}^{\infty} B_i \implies a \in A_j$, for all j. Thus, by construction, for all j,

$$\frac{3}{4} r < \frac{3}{4} \sup\{r \ ; \ B \in \mathcal{F}, a \in A_j\} \leq r_j$$

By the claim 3, $r_j \to 0$. which contradicts the non-degeneracy of the balls, that is $r \neq 0$. This proves the claim 4.

Fix $k > 1$ and set

$$I := \{j \ ; \ 1 \leq j \leq k, B_j \cap B_k \neq \phi\}$$

$$K := I \cap \{j \ ; \ r_j \leq 3r_k\}$$

$$I \equiv K \cup (I - K)$$

Claim 5 : Cardinality of K constructed above is less than or equal to 20^n (which is independent of k).
Proof of the claim 5 : Let $j \in K$. Then,by the construction of K, $B_j \cap B_k \neq \phi$ and

$r_j \leq 3r_k$. Choose any $x \in B(a_j, \frac{r_j}{3})$. Then,as $B_j \cap B_k \neq \phi$,

$$|x - a_k| \leq |x - a_j| + |a_j - a_k| \leq \frac{r_j}{3} + r_j + r_k = \frac{4}{3}r_j + r_k \leq 4r_k + r_k = 5r_k$$

Thus

$$B(a_j, \frac{r_j}{3}) \subset B(a_k, 5r_k)$$

Claim 2 tells that the balls with 1/3rd radius are disjoint and hence integrating, we get

$$\alpha(n)5^n r_k^n = \mathcal{L}^n\big((B(a_k, 5r_k)\big) \geq \sum_{j \in K} \mathcal{L}^n\big(B(a_j, \frac{r_j}{3})\big) = \sum_{j \in K} \alpha(n) \left(\frac{r_j}{3}\right)^n \cdots$$

By the claim 1 and noting that $j < k$ in the definiton of K,

$$\cdots \geq \sum_{j \in K} \alpha(n) \left(\frac{r_k}{4}\right)^n = Card(K)\alpha(n) \left(\frac{r_k^n}{4^n}\right)$$

This proves the claim 5.

Estimate on $Card(I - K)$:
Let $i, j \in I - K$ with $i \neq j$.
Then

- $1 \leq i, j < k$.
- $B_i \cap B_k \neq \phi$, $3r_k < r_i$.
- $B_j \cap B_k \neq \phi$, $3r_k < r_j$.

For simplicity, let $a_k = 0$ and a_i, a_j be seen as the vectors with magnitude $|a_i - a_k|$ and $|a_j - a_k|$ respectively from 0 and preserving the direction. Let $0 \leq \theta \leq 2\pi$ be the angle between a_i and a_j.
The idea is to find the lower bound on θ.
Some observations to be noted :
Since $i, j < k$ and $0 = a_k \notin B_i \cup B_j$,

- $r_i < |a_i|$ and $r_j < |a_j|$.
- $B_i \cap B_k \neq \phi$ and $B_j \cap B_k \neq \phi \implies |a_i| \leq r_i + r_k$ and $|a_j| \leq r_j + r_k$.
- W.L.O.G , assume $|a_i| \leq |a_j|$.

Putting this in a compact form, we have the following:

$$3r_k < r_i < |a_i| \le r_i + r_k$$

$$3r_k < r_j < |a_j| \le r_j + r_k$$

$$|a_i| \le |a_j|$$

We prove that there is a constant $\theta_0 > 0$ such that it depends only on the dimension of the space and for $i \neq j$, $1 \le i, j \le k$ with $B_i \cap B_k \neq \phi$, $B_j \cap B_k \neq \phi$, the angle between $a_i - a_k$ and $a_j - a_k$ is greater than θ_0.

Claim 6 : $\{$ If $\cos(\theta) > \frac{5}{6}$, then $a_i \in B_j\} \equiv \{a_i \notin B_j \implies \cos(\theta) \le \frac{5}{6}\}$.
Proof of the claim 6 :
Case 1 : $|a_i - a_j| > |a_j|$
The law of cosines gives

$$\cos(\theta) = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|}$$

$$\le \frac{|a_i|^2}{2|a_i||a_j|} = \frac{|a_i|}{2|a_j|} \le \frac{1}{2} < \frac{5}{6}$$

Thus $a_i \notin B_j$ and $|a_i - a_j|$ is greater than the radius of B_j. The contra-positive statement requires that $\cos(\theta) \le \frac{5}{6}$ which is got by the law of cosines as above.
Case 2 : If $a_i \notin B_j$ and $|a_i - a_j| \le |a_j|$. Then, as $r_j < |a_i - a_j|$, we have

$$cos(\theta) = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|}$$

$$= \frac{|a_i|}{2|a_j|} + \frac{(|a_j| - |a_i - a_j|)(|a_j| + |a_i - a_j|)}{2|a_i||a_j|}$$

$$\le \frac{1}{2} + \frac{(|a_j| - |a_i - a_j|)(2|a_j|)}{2|a_i||a_j|}$$

$$\le \frac{1}{2} + \frac{r_j + r_k - r_j}{r_i}$$

$$= \frac{1}{2} + \frac{r_k}{r_i} < \frac{1}{2} + \frac{1}{3} \le \frac{5}{6}$$

This proves the claim 6.

Claim 7 : If $a_i \in B_j$, then

$$0 \leq |a_i - a_j| + |a_i| - |a_j| \leq |a_j| g(\theta)$$

where

$$g(\theta) = \frac{8}{3}\big(1 - \cos(\theta)\big)$$

Proof of the claim 7 : Since $a_i \in B_j$, we must have $i < j$ and hence $a_j \notin B_i$ and so $|a_i - a_j| > r_i$. Thus

$$0 \leq \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|}$$

As $|a_i| \leq |a_j| \implies |a_i - a_j| - |a_i| + |a_j| \geq |a_i - a_j| \implies \frac{|a_i - a_j| - |a_i| + |a_j|}{|a_i - a_j|} \geq 1$,

$$\frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} \leq \frac{|a_i - a_j| + |a_i| - |a_j|}{|a_j|} . \frac{|a_i - a_j| - |a_i| + |a_j|}{|a_i - a_j|}$$

$$= \frac{|a_i - a_j|^2 - (|a_i| - |a_j|)^2}{|a_j||a_i - a_j|}$$

$$= \frac{|a_i|^2 + |a_j|^2 - 2|a_i||a_j|\cos(\theta) - |a_i|^2 - |a_j|^2 + 2|a_i||a_j|}{|a_j||a_i - a_j|}$$

$$= \frac{2|a_i|(1 - \cos(\theta))}{|a_i - a_j|}$$

$$\leq \frac{2(r_i + r_k)(1 - \cos(\theta))}{r_i} \leq \frac{2(1 + \frac{1}{3})r_i(1 - \cos(\theta))}{r_i} \equiv g(\theta)$$

This proves the claim 7.

Claim 8 : If $a_i \in B_j$, then $\cos(\theta) \leq \frac{61}{64}$.
Proof of the claim 8 : Since $a_i \in B_j$ and $a_j \notin B_i$, we have $r_i < |a_i - a_j| \leq r_j$. Since $i < j$ and $r_j \leq \frac{4}{3} r_i$,

$$\begin{aligned}
|a_i - a_j| + |a_i| - |a_j| &\geq r_i + r_i - r_j - r_k \\
&\geq \frac{3}{2} r_j - r_j - r_k \\
&= \frac{1}{2} r_j - r_k \\
&\geq \frac{1}{6} r_j \\
&= \frac{1}{6} . \frac{3}{4} . \left(r_j + \frac{1}{3} r_j \right) \\
&\geq \frac{1}{8}(r_j + r_k) \\
&\geq \frac{1}{8}|a_j|
\end{aligned}$$

Thus by the previous claim,

$$\frac{1}{8}|a_j| \leq |a_i - a_j| + |a_i| - |a_j| \leq |a_j| g(\theta)$$

Hence ,

$$\cos(\theta) \leq \frac{61}{64}$$

This proves the claim 8.

Claim 9 : For all $i, j \in I - K$ with $i \neq j$, let θ be the angle between $a_i - a_k$ and $a_j - a_k$. Then,

$$\theta \geq \cos^{-1}\left(\frac{61}{64}\right) \equiv \theta_0 > 0$$

Proof of the claim 9 : From the claim 6, 7 and 8,

$$\cos(\theta) \leq \frac{61}{64} \text{ or } \cos(\theta) \leq \frac{5}{6}$$

Since, $\frac{5}{6} < \frac{61}{64}$, and observe that cos is a decreasing function in $[0, \frac{\pi}{2}]$,

$$0 < \theta_0 \equiv \cos^{-1}\left(\frac{61}{64}\right) \leq \theta \leq \cos^{-1}\left(\frac{5}{6}\right)$$

This proves the claim 9.

Claim 10: There is a constant L_n depending only on the dimension of the space such that $Card(I - K) \leq L_n$.
Proof of the claim 10: Fix $r_0 > 0$ such that if $x_0 \in \partial B(0, 1)$ and for $y, z \in B(x_0, r_0)$, the angle between y and z from the origin is less than θ_0, where θ_0 is defined in claim 9. The idea involved is considering sectors of the circles dictated on the boundary as above. By compactness of the boundary, choose L_n so that $\partial B(0, 1)$ can be covered by L_n many balls with radius r_0 and centers lying on $\partial B(0, 1)$, but cannot be covered by L_{n-1} balls.
That is, for any y, z in one of the L_n balls,

$$\frac{< y, z >}{|y||z|} < \cos(\theta_0) \equiv \frac{61}{64}$$

Consider $B(0, t)$ and $x_0 \in \partial B(0, t)$.

$$y \in B(x_0, tr_0) \iff \frac{y}{t} \in B\big(\frac{x_0}{t}, r_0\big) \text{ with } \frac{x_0}{t} \in \partial B(0, 1)$$

Therefore, $\partial B(0, 1)$ is covered by L_n (not L_{n-1}) many balls of radius r_0 if and only if $\partial B(0, t)$ is covered by L_n (not L_{n-1}) many balls of radius tr_0. Now, with the same logic,

for the above fixed k, $B(x_k, r_k)$, treating $x_k \equiv 0$ and $r_k \equiv t$, we have $\partial B(x_k, r_k)$ can be covered by L_n many balls of radius $r_k r_0$ and not by L_{n-1} many balls of radius $r_k r_0$. Also, by the claim 9, if $i, j \in I - K$ and $i \neq j$, the angle between $a_i - a_k$ and $a_j - a_k$ is greater than or equal to θ_0 and ∂B_k is covered by L_n (not L_{n-1}) many balls implies the rays $a_i - a_k$ and $a_j - a_k$ cannot go through the same ball on ∂B_k.
Thus $Card(I - K) \leq L_n$, which is independent of k. This proves the claim 10.

Now , set

$$M_n := 20^n + L_n + 1$$

Thus, clearly

$$Card(I) < M_n$$

Claim 11 : Let

$$\mathcal{G} := \{B_i \equiv B(a_i, r_i) \; ; \; 1 \leq i \leq J\}$$

For $k \geq M_n + 1$ and define $\mathcal{G}_{ik}$ for $1 \leq i \leq M_n$ as :

- $B_i \in \mathcal{G}_{ik}$; $\forall k$, $1 \leq i \leq M_n$.
- $\mathcal{G}_{ik}$ consists of disjoint balls from $\mathcal{G}$ such that if $B \in \mathcal{G}_{ik}$ and $B \neq B_i$, then there is some j such that $M_n + 1 \leq j \leq k$ such that $B = B_j$.
- If $B \in \mathcal{G}_{ik}$ then $B \notin \mathcal{G}_{jk}$, for all $j \leq i - 1$.

Let

$$\mathcal{G}_i := \bigcup_{k=M_n+1}^{\infty} \mathcal{G}_{ik}$$

$$\text{Then, } A \subset \bigcup_{B \in \mathcal{G}} B = \bigcup_{i=1}^{M_n} \Big(\bigcup_{B \in \mathcal{G}_i} B \Big)$$

Proof of the claim 11 : (Existence of $\{\mathcal{G}_{ij}\}$.) The proof for the existence of the family of sets follows by induction on k.
Declare, for $1 \leq i \leq M_n$ and for any $k > M_n$ that

$$\mathcal{G}_{ik} \text{ consists of } B_i \text{ and hence non empty.}$$

For $k = M_n + 1$, by the previous calculation, the estimate on the cardinality of I gives

$$Card\{1 \leq i \leq k \equiv M_n + 1 \; ; \; B_i \cap B_k \neq \phi\} < M_n$$

Thus, by the strict inequality as above, there is an $i \leq M_n$ such that $B_k \cap B_i = \phi$. Choose a minimal i such that $B_k \cap B_i = \phi$ and $B_k \cap B_j \neq \phi, \forall j \leq i - 1$. Hence $B_k \in \mathcal{G}_{ik}$.
Now, assume the statement is true upto k (which is greater than $M_n + 1$).

As $Card\{i \leq k\ ;\ B_i \cap B_{k+1} \neq \phi\} \leq M_n$, there is one $1 \leq i \leq M_n$ such that $B_{k+1} \cap B_i = \phi$, for all $B \in \mathcal{G}_{ik}$. Hence, by taking the minimum i such that the above happens,

$$\implies B_{k+1} \in \mathcal{G}_{i,k+1}$$

This proves the existence of $\{\mathcal{G}_{ij}\}$ and by claim 4,

$$A \subset \bigcup_{B \in \mathcal{G}} B \equiv \bigcup_{i=1}^{M_n} \Big(\bigcup_{B \in \mathcal{G}_i} B \Big)$$

This proves the claim 11.

We proved the result under the assumption that A is bounded.
Relaxing that condition and assume A is unbounded. Partition A as follows:
For $l \geq 1$, declare,

$$A_l := A \cap \{x\ ;\ 3D(l-1) \leq |x| < 3Dl\}$$

$$\mathcal{F}_l := \{B(a,r) \in \mathcal{F}\ ;\ a \in A_l\}$$

Since A_l is bounded, by the previous step, there is finite family of countable collection of disjoint closed balls,say $\mathcal{G}_1^l, \mathcal{G}_2^l, \ldots, \mathcal{G}_{M_n}^l$ in $\mathcal{F}_l$ such that

$$A_l \subset \bigcup_{i=1}^{\infty} \bigcup_{B \in \mathcal{G}_i^l} B$$

Declare

$$\mathcal{G}_j^1 = \bigcup_{l=1}^{\infty} \mathcal{G}_j^{2l-1} \text{ for } 1 \leq j \leq M_n$$

$$\mathcal{G}_j^2 = \bigcup_{l=1}^{\infty} \mathcal{G}_j^{2l} \text{ for } 1 \leq j \leq M_n$$

$$\implies A \subset \left[\bigcup_{j=1}^{M_n} \Big(\bigcup_{B \in \mathcal{G}_j^1} B \Big) \right] \bigcup \left[\bigcup_{j=1}^{M_n} \Big(\bigcup_{B \in \mathcal{G}_j^2} B \Big) \right]$$

Re-labelling the balls, we get that the required dimensional constant for the theorem N is $2M_n$ and the family $\mathcal{G}_i$ consists of countable disjoint closed balls. This proves the theorem.

$\square$

Corollary 2.6. *Let m be a borel, regular measure on $\mathbb{R}^n$ and $\mathcal{F}$ be any collection of non degenerate closed balls. Let A denote the set of centers of the balls in $\mathcal{F}$. Assume $m(A) < \infty$ and*

$$\inf\{r\ ;\ B(a,r) \in \mathcal{F}\} = 0 \textit{ for each } a \in A$$

Then, for each open set $U \subset \mathbb{R}^n$, there is a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{G}} B \subset U$$

and

$$m\left((A \cap U) - \bigcup_{B \in \mathcal{G}} B\right) = 0$$

Remark 2.7. Here, A need not be $m-$ measurable.

Proof. Fix $1 - \frac{1}{N} < \theta < 1$, where N is the dimensional constant as in the Besicovitch theorem.

Claim : There is a set of finite disjoint closed balls $B_1, B_2, \ldots, B_{M_1}$ in U such that

$$m\left((A \cap U) - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta m(A \cap U)$$

Proof of The claim : Let

$$\mathcal{F}_1 := \{B \ ; \ B \in \mathcal{F}, diam B \leq 1, B \subset U\}$$

Note that the centers of the balls in $\mathcal{F}_1$ are precisely $A \cap U$. This is true as

- Clearly $B \subset U$. By the defintion of $\mathcal{F}$, $B \subset A \cap U$
- Let $a \in A \cap U$. Then, there is r, such that $0 < r < 1$ and $B(a, r) \subset U$ and $B(a, r) \in \mathcal{F}$. This is by the infimum property mentioned in the hypothesis. Thus a is centre of some ball in $\mathcal{F}_1$.

Thus, by the Besicovitch theorem, there are families of disjoint balls $\mathcal{G}_1, \mathcal{G}_2, \ldots \mathcal{G}_N$ from $\mathcal{F}_1$ such that

$$(\mathcal{A} \cap U) \subset \bigcup_{i=1}^{N} \bigcup_{B \in \mathcal{G}_i} B$$

And thus,

$$m(A \cap U) \leq \sum_{i=1}^{N} m\left((A \cap U) \cap (\bigcup_{B \in \mathcal{G}_i} B)\right)$$

Thus there is a j such that

- $1 \leq j \leq N$
- $m\Big((A \cap U) \cap \bigcup_{B \in \mathcal{G}_j} B\Big) \geq \frac{1}{N} m(A \cap U)$

If not, there $m\left((A \cap U) \cap (\bigcup_{B \in \mathcal{G}_i} B)\right) < \frac{1}{N} m(A \cap U)$, for all i and thus summing it,

$$m(A \cap U) \leq \sum_{i=1}^{N} m\left((A \cap U) \cap (\bigcup_{B \in \mathcal{G}_i} B)\right) < \frac{N}{N} m(A \cap U) \implies m(A \cap U) < m(A \cap U)$$

Now, as per theorem 1.43 there are balls $B_1, B_2, \ldots B_{M_1} \in \mathcal{G}_j$ such that

$$m\left((A \cap U) \cap (\bigcup_{i=1}^{M_1} B)\right) \geq (1-\theta) m(A \cap U)$$

Now $\bigcup B_i$ is measurable, implies

$$m(A \cap U) = m\left((A \cap U) \cap \bigcup_{i=1}^{M_1} B\right) + m\left((A \cap U) - \bigcup_{i=1}^{M_1} B\right)$$

Thus ,

$$m\left((A \cap U) - \bigcup_{i=1}^{M_1} B_i\right) \leq \theta m(A \cap U)$$

Now, Declare

$$U_2 := U - \bigcup_{i=1}^{M_1} B_i$$

$$\mathcal{F}_2 := \{B;\ B \in \mathcal{F}, diam B \leq 1, B \subset U_2\}$$

Clearly, the centers of $\mathcal{F}_2$ are precisely $A \cap U_2$, by the same logic as before. Hence, There are disjoint balls $B_{M_1+1}, \ldots, B_{M_2} \in \mathcal{F}_2$ such that

$$m\left((A \cap U) - \bigcup_{i=1}^{M_2} B_i\right) = m\left((A \cap U_2) - \bigcup_{i=M_1+1}^{M_2} B_i\right)$$

$$\leq \theta m(A \cap U_2) \leq \theta^2 m(A \cap U)$$

This process gives a countable collection of disjoint balls from $\mathcal{F}$, within U such that

$$m\left((A \cap U) - \bigcup_{i=1}^{M_k} B_i\right) \leq \theta^k m(A \cap U)$$

$\theta^k \to 0$ as $k \to \infty$ concludes the corollary. $\square$

2.2 Differentiation of Radon measures.

Definition 2.8. Let u, v be two radon measures on $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, define

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

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

Definition 2.9. If $\overline{D}_u v(x) = \underline{D}_u v(x) < +\infty$, we say that v is differentiable with respect to u at x and denote

$$D_u v(x) \equiv \overline{D}_u v(x) = \underline{D}_u v(x)$$

$D_u v$ is called the derivative of v with respect to u. It is also called as density of v with respect to u.

Lemma 2.10. *Fix $0 < \alpha < \infty$. Then*

- $A \subset \{x \in \mathbb{R}^n \ ; \ \underline{D}_u v(x) \le \alpha\} \implies v(A) \le \alpha u(A) \ldots\ldots\ldots (i)$
- $A \subset \{x \in \mathbb{R}^n \ ; \ \overline{D}_u v(x) \ge \alpha\} \implies v(A) \ge \alpha u(A) \ldots\ldots\ldots (ii)$

Proof. Assume $u(\mathbb{R}^n)$ and $v(\mathbb{R}^n)$ are finite. Fix $\epsilon > 0$. Let U be open and $A \subset U$, where A satisfies the hypothesis of (i) . Declare

$$\mathcal{F} := \{B \ ; \ B \equiv B(x,r), x \in A, B \subset U, v(B) \le (\alpha + \epsilon)u(B)\}$$

Then

$$\text{Claim : For all } x \in A, \ \ \inf\{r \ ; \ B(x,r) \in \mathcal{F}\} = 0$$

Proof of the claim : Notice that $A \subset \{x \in \mathbb{R}^n \ ; \ \underline{D}_u v(x) \le \alpha\}$

$$\implies \liminf_{r\to 0} \frac{v(B(x,r))}{u(B(x,r))} \le \alpha \ , \text{ for } x \in A$$

$$\implies \lim_{\epsilon'\to 0} \inf_{|r|<\epsilon'} \frac{v(B(x,r))}{u(B(x,r))} \le \alpha$$

There is $\epsilon' > 0$, small enough such that

$$\inf_{|r|<\epsilon''} \frac{v(B(x,r))}{u(B(x,r))} < \alpha + \frac{\epsilon}{2} \ ; \ \forall \epsilon'' < \epsilon'$$

and

As U is open, $B(x,r) \subset U$, for all $|r| < \epsilon'$

Thus there is $r' < \epsilon'$ such that

$$\frac{v(B(x,r'))}{u(B(x,r'))} < \alpha + \epsilon$$

$$\implies v(B(x,r')) < u(B(x,r'))(\alpha + \epsilon)$$

This is true for all $\epsilon'' < \epsilon'$. Thus sending ϵ'' to 0 and hence, r to 0, concludes the claim. This proves the claim.
By the corollary 2.6 (the corollary to the besicoivitch theorem), there is countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$v\left(A - \bigcup_{B\in\mathcal{G}} B\right) = 0$$

Since $\bigcup_{B\in\mathcal{G}} B$ is measurable,

$$\begin{aligned} v(A) = v\left(A \cap (\bigcup_{B\in\mathcal{G}} B)\right) + v\left(A - \bigcup_{B\in\mathcal{G}} B\right) \\ \leq \sum_{B\in\mathcal{G}} v(B) \\ \leq \sum_{B\in\mathcal{G}} (\alpha+\epsilon)u(B) \\ \leq (\alpha+\epsilon)u(U) \end{aligned}$$

This estimate is valid for each U such that $A \subset U$.
Regularity result of the radon measures says that any measurable set can be approximated by the open sets from the outside. That is, for m to be radon measure, we have

$$m(A) = \inf \{ \ m(U) \ ; \ A \subset U, \ U \text{ is open } \}$$

Thus, taking infimum on the estimate obtains,

$$v(A) \leq (\alpha+\epsilon)u(A)$$

Noting that $\epsilon > 0$ was arbitrary, (i) is proved, for the case of finite measure on the whole space, as per the assumption.
For the non finite case, let $U_k \equiv B(0,k) \subset \overline{B(0,k)}$.

$$\implies v(U_k) < \infty \ , \ v(U_k) < \infty$$

Declare

$$u_k := (u|U_k)$$

$$v_k := (v|U_k)$$

Clearly, u_k and v_k are finite radon measures and now, let $A_k = A \cap U_k$.
Since, U_k's are open and

$$A_k \subset \{x \in U_k \ ; \ \underline{D_{u_k}} v_k(x) \leq \alpha\} \ ; \ A_k \subset A_{k+1} \ ; \ A = \bigcup_{k=1}^{\infty} A_k$$

$$\implies v(A \cap U_k) = v_k(A_k) \leq \alpha u_k(A_k) = \alpha u(A \cap U_k) \leq \alpha u(A)$$

Let $k \to \infty$ to get

$$v(A) = \lim_{k\to\infty} v\big(A \cap V_k\big) \leq \alpha u(A)$$

This proves (i) of the lemma.
(ii) of the lemma has the same steps as for the first part and the proof follows.

□

Definition : Let X be a set and Y be a topological set. Assume that m is (an outer) measure on X.
A function $f : X \to Y$ is said to be $m-$ measurable if for every open set $U \subset Y$, $f^{-1}(U)$ is $m-$ measurable.

Theorem 2.11. *Let u, v be radon measures on $\mathbb{R}^n$. Then $D_u v$ exists and is finite u almost everywhere. Furthermore, $D_u v$ is $u-$ measurable.*

Proof. Assume that $v(\mathbb{R}^n)$ and $u(\mathbb{R}^n)$ are finite.
Claim 1 : $D_u v$ exists and is finite u almost everywhere.
Proof of the claim 1 : Let

$$I := \{x \ ; \ \underline{D}_u v(x) = +\infty\}$$

and for all $0 < a < b$, let

$$R(a, b) := \{x \ ; \ \underline{D}_u v(x) < a < b < \overline{D}_u v(x) < \infty\}$$

For each $\alpha > 0$, $I \subset \{x \ ; \ \overline{D}_u v(x) \geq \alpha\}$. By the previous lemma,

$$u(I) \leq \frac{1}{\alpha} v(I)$$

Let $\alpha \to \infty$, we see that $u(I) = 0$ and so $\overline{D}_u v$ is finite u almost everywhere.
Now,

$$R(a, b) \subset \{x \ ; \ \overline{D}_u v(x) \geq b\}$$

and

$$R(a,b) \subset \{x \ ; \ \underline{D}_u v(x) \leq a\}$$

Thus,by the previous lemma,

$$b \ u(R(a,b)) \leq v(R(a,b)) \leq a \ u(R(a,b))$$

Since $b > a$, we have $u(R(a,b)) = 0$. Also,

$$\{x \ ; \ \underline{D}_u v(x) < \overline{D}_u v(x) < \infty\} = \bigcup_{0<a<b \ ; \ a,b \in \mathbb{Q}} R(a,b)$$

As a consequence, $D_u v$ exists and is finite u almost everywhere. This proves the claim 1.

Claim 2 : For each $x \in \mathbb{R}^n$ and $r > 0$,

$$\limsup_{y \to x} u(B(y,r)) \leq u(B(x,r))$$

$$\limsup_{y \to x} v(B(y,r)) \leq v(B(x,r))$$

Proof of the claim 2 : Choose $\{y_k\}$ in $\mathbb{R}^n$ such that $y_k \to x$. Set

$$f_k := \mathcal{X}_{B(y_k,r)}$$

$$f := \mathcal{X}_{B(x,r)}$$

Then,

$$(S) : \limsup_{k \to \infty} f_k \leq f$$

This is true as :

Case 1 : $f \equiv 1$. Since $f_k = 0$ or 1, the statement (S) is true for this case.

Case 2 : If $f(y) = 0$, i.e y is not in $B(x,r)$. Say $y \notin \partial B(x,r)$. Clearly, $|x - y| > r$ and there is a $k_0 > 0$ such that for all $k \geq k_0$ we have that $|y_k - x| < \min\{|y - x|, \frac{r}{2}\}$ and thus $B(y_k, r) \cap \{y\} = \phi$. Hence, $f_k(y) = 0$ for all $k \geq k_0$. This proves the statement (S) for this special case.

Case 3 : Same conditions as in case 2, but, if y is on the boundary. Note that, as per the assumption made earlier, the balls are closed if nothing is mentioned about them. So, $y \notin B(x,r)$ implies $y \notin \partial B(x,r)$ and hence, only case 2 is possible.

Thus, combining all the 3 cases, the statement (S) is true.

Now,

$$\liminf_{k \to \infty} (-f_k) \geq -f$$

$$\implies \liminf_{k \to \infty} (1 - f_k) \geq 1 - f$$

Thus, by the fatou's lemma,

$$\int_{B(x,2r)}(1-f)du \leq \int_{(B(x,2r)} \liminf_{k\to\infty}(1-f_k)du \leq \liminf_{k\to\infty}\int_{B(x,2r)}(1-f_k)du$$

Hence,

$$u(B(x,2r)) - u(B(x,r)) \leq \liminf_{k\to\infty}\Big[u(B(x,2r)) - u(B(y_k,2r))\Big]$$

Now, as u is a radon measure and $\overline{B(x,2r)}$ is compact implies

$$u(B(x,2r)) \leq u(\overline{B(x,2r)}) < \infty$$

Thus

$$\limsup_{y\to x} u(B(y,r)) \leq u(B(x,r))$$

The same result holds for v as well and this proves claim 2.

Claim 3 : $D_u v$ is measurable with respect to u.
Proof of the claim 3 : By the claim 2,

$$x \to u(B(x,r))$$

$$x \to v(B(x,r))$$

are upper semi-continuous and thus borel measurable. As a consequence, for every $r > 0$,

$$f_r(x) = \begin{Bmatrix} \frac{v(B(x,r))}{u(B(x,r))} & \text{if } u(B(x,r)) > 0 \\ +\infty & \text{if } u(B(x,r)) = 0 \end{Bmatrix}$$

f_r defined above is $u-$ measurable.
Note that

$$D_u v = \lim_{r\to 0} f_r = \lim_{k\to\infty} f_{1/k}$$

The above is true for u almost everywhere as per the claim 1. So $D_u v$ is u measurable. This proves the claim 3.

Suppose that u, v are not finite, observing that u, v are radon, we have, for $k \in \mathbb{N}$,

$$u\big(B(0,k)\big) \leq u\left(\overline{B(0,k)}\right) < \infty$$

$$v\big(B(0,k)\big) \leq v\left(\overline{B(0,k)}\right) < \infty$$

As before, declare

$$u_k := \big(u|B(0,k)\big)$$

$$v_k := \big(v|B(0,k)\big)$$

Thus, for $x \in B(0,k)$,

$$\overline{D_{u_k} v_k}(x) = D_u v(x)$$

$$\underline{D_{u_k} v_k}(x) = D_u v(x)$$

Hence, if $I_k := I \cap B(0,k)$, then $u(I_k) = u_k(I_k) = 0 \implies u(I) = \lim_{k\to\infty} u(I_k) = 0$. Similarly, if $a < b$, let

$$R_k(a,b) := \{x \in B(0,k) \ ; \ \underline{D_{u_k} v_k}(x) < a < b < \overline{D_{u_k} v_k}(x)\}$$

$$\implies R_k(a,b) \subset B(0,k) \ , \ u_k\big(R_k(a,b)\big) = u\big(R_k(a,b)\big) = 0$$

$$\implies u\big(R(a,b)\big) = \lim_{k\to\infty} u\big(R_k(a,b)\big) = 0$$

Hence, $\overline{D}_u v(x) = \underline{D}_u v(x)$ exist for $u-$ almost everywhere and $x \to D_u v(x)$ is a u_k-measurable function for all k implies for an open set $U \subset \mathbb{R}^n$, $\big(D_u v^{-1}(U)\big)$ is u_k- measurable for all k. That is, for any set $B \subset \mathbb{R}^n$,

$$u_k(B) = u_k\big(B \cap D_u v^{-1}(U)\big) + u_k\big(B - D_u v^{-1}(U)\big)$$

The above is true for all k and hence, taking the limit $k \to \infty$, we see that $x \to D_u v(x)$ is u measurable.

$\square$

2.2.1 Radon-Nikodym Derivative

Definition 2.12. The measure v is absolutely continuous with respect to u, written as

$$v \ll u$$

if

For all $A \subset \mathbb{R}^n$, with $u(A) = 0$, then $v(A) = 0$

Definition 2.13. The measures u and v are mutually singular, denoted as

$$v \perp u$$

if

If there is a $B \subset \mathbb{R}^n$, borel, such that $u(\mathbb{R}^n - B) = v(B) = 0$

Theorem 2.14. ***Radon-Nikodym Theorem/ Differentiation theorem for Radon Measures***
Let u, v be radon measures on $\mathbb{R}^n$ with $v \ll u$. Then, for all A which are $u-$ measurable subsets of $\mathbb{R}^n$,

$$v(A) = \int_A D_u v \; du$$

Proof. Let A be $u-$ measurable. This implies there is a borel set B such that $u(A) = u(B)$.

$$\implies u(A - B) = 0$$

$$\implies v(A - B) = 0$$

This tells that A is also $v-$measurable. Set

$$Z := \{x \in \mathbb{R}^n \; ; \; D_u v(x) = 0\}$$

$$I := \{x \in \mathbb{R}^n \; ; \; D_u v(x) = +\infty\}$$

Let

$$Z_k := Z \cap B(0, k) \; , \; k \in \mathbb{N}$$

Then, for any $\alpha > 0$, we have $Z_k \subset \{x \in B(0, k) \; ; \; D_u v(x) \leq \alpha\}$.
From the previous lemma, we therefore have $v(Z_k) \leq \alpha u(Z_k)$. Letting $\alpha \to 0$,

$$\implies v(Z) = \lim_{k \to 0} v(Z_k) = 0$$

Note that this also gives that Z is v measurable, as v is a complete measure.

Case 1 : The case for Z.

$$v(Z) = 0$$

$$\int_Z D_u v \; du = \int_Z 0 \; du = 0$$

This tells that the theorem is true for Z.
Case 2 : The case for I. It is noted that $v(I) = 0$. And by the previous theorem, $D_u v$

is finite $u-$almost everwhere and hence,

$$\int_I D_u v \ du = 0$$

This proves the theorem for I.

Case 3: For any other $u-$measurable set A. Choose and fix $1 < t < \infty$. Define , for $m \in \mathbb{Z}$,

$$A_m := A \cap \{x \in \mathbb{R}^n \ ; \ t^m \leq D_u v(x) < t^{m+1}\}$$

A is $u-$ measurable and $\{x \in \mathbb{R}^n \ ; \ t^m \leq D_u v(x) < t^{m+1}\} = D_u v^{-1}\left([t^m, t^{m+1})\right)$ is also $u-$ measurable as per the previous theorem and hence, A_m is $u-$ measurable and thus $v-$ measurable.

Also

$$A - \bigcup_{-\infty}^{+\infty} A_m \subset \left(Z \cup I \cup \{x \ ; \ \overline{D}_u v(x) \neq \underline{D}_u v(x)\}\right)$$

Thus,

$$v\left(A - \bigcup_{-\infty}^{+\infty} A_m\right) = 0$$

As a consequence,

$$v(A) = \sum_{-\infty}^{+\infty} v(A_m)$$

$$\leq \sum_{-\infty}^{+\infty} t^{m+1} u(A_m)$$

$$= t \sum_{-\infty}^{+\infty} t^m u(A_m)$$

$$\leq t \sum_m \int_{A_m} D_u v \ du$$

As, A_i's are disjoint,

$$= t \int_A D_u v du$$

Also

$$v(A) = \sum_{m=-\infty}^{+\infty} v(A_m)$$

$$\geq \sum_{m=-\infty}^{+\infty} t^m v(A_m)$$

$$= \frac{1}{t} \sum_{m=-\infty}^{+\infty} t^{m+1} u(A_m)$$

$$\geq \frac{1}{t} \sum_{m=-\infty}^{+\infty} t^{m+1} u(A_m)$$

$$\geq \frac{1}{t} \sum_{m=-\infty}^{\infty} \int_{A_m} D_u v \ du$$

As, A_i's are disjoint,

$$= \frac{1}{t} \int_A D_u v \ du$$

Hence,

$$\frac{1}{t} \int_A D_u v \ du \leq v(A) \leq t \int_A D_u v \ du$$

Notice that t was arbitrarily greater than 1. Thus taking $t \to 1+$, we can conclude that

$$v(A) = \int_A D_u v \ du$$

This proves the thoerem.

□

Theorem 2.15. ***Lebesgue Decomposition theorem :***
Let u, v be radon measures on $\mathbb{R}^n$.

- *Then*

$$v = v_{ac} + v_s$$

$$v_{ac} \ll u$$

$$v_s \perp u$$

- *Furthermore, for $u-$ almost everywhere,*

$$D_u v = D_u v_{ac}$$

$$D_u v_s = 0$$

and

$$v(A) = \int_A D_u v \ du + v_s(A)$$

Definition 2.16. v_{ac} is the absolutely continuous part of v with respect to u.
v_s is the singular of v with respect to u.

Proof. Assume $v(\mathbb{R}^n)$ is finite.
Let

$$\epsilon := \{A \subset \mathbb{R}^n \ ; \ \ A \text{ is borel }, u(\mathbb{R}^n - A) = 0\}$$

Clearly, ϵ is not empty as $\mathbb{R}^n \in \epsilon$. By the infimum property, choose $B_k \in \epsilon$ such that

$$v(B_k) \leq \inf_{A\in\epsilon} v(A) + \frac{1}{k}$$

Set

$$B = \bigcap_{k=1}^{\infty} B_k$$

Since,

$$u(\mathbb{R}^n - B) \leq \sum_{k=1}^{\infty} u(\mathbb{R}^n - B_k) = 0$$

$$\implies u(\mathbb{R}^n - B) = 0 \implies B \in \epsilon$$

So,

$$v(B) = \inf_{A\in\epsilon} v(A)$$

Declare

$$v_{ac} := (v|B)$$

$$v_s := (v|(\mathbb{R}^n - B))$$

Clearly, $v_s(B) = 0, u(\mathbb{R}^n - B) = 0$ and thus $v_s \perp u$.
Let A be a borel set such that $u(A) = 0$.
Case 1 : If $A \cap B = \phi$, then clearly, $v_{ac}(B) = 0$.
Case 2 : If $A \subset B$ with $u(A) = 0$.
Claim : $v(A) = 0$.
Suppose not, say $v(A) > 0$. Since A is a borel set

$$\implies B - A \in \epsilon \text{ , as } u\big(\mathbb{R}^n - (B - A)\big) = u\big((\mathbb{R}^n - B) \cup A\big) = u\big(\mathbb{R}^n - B\big) + u(A) = 0.$$

$$\implies \inf_{C\in\epsilon} v(C) \leq v(B - A) = v(B) - v(A) < v(B) = \inf_{C\in\epsilon} v(C)$$

This is a contradiction and thus proves the claim.
If $A \subset B$ and not borel, but with $u(A) = 0$, then, by the borel regularity, choose A_1, a borel set such that

- $A \subset A_1$
- $u(A_1) = 0$

Declare $A_2 := A_1 \cap B$.

$$\implies A \subset A_2 \subset B$$

A_2 is a borel set and $u(A_2) = 0$. Hence, by the claim, $v(A_2) = 0$ and therefore, $v_{ac}(A) = v(A) = 0$. This proves that $v_{ac} \ll u$ and $v_{ac} = v$ on every subset of B.

Finally, fix $\alpha > 0$ and set

$$C := \{x \in B\ ;\ D_u v_s(x) \geq \alpha\} \subset B$$

By the earlier lemma 2.10,

$$\alpha\ u(C) \leq v_s(C) = 0$$

$$\implies u(C) = 0 \text{ as } \alpha > 0$$

Thus $D_u v_s \equiv 0$, $u-$ almost everywhere. Now,

$$D_u v(x) = \begin{cases} \lim_{r\to 0} \frac{v(B(x,r))}{u(B(x,r))} & \text{if, for all } r > 0,\ u(B(x,r)) > 0 \\ +\infty & \text{else.} \end{cases}$$

$v = v_{ac} + v_s$ implies,

$$D_u v(x) = \begin{cases} \lim_{r\to 0} \frac{v_{ac}(B(x,r))}{u(B(x,r))} + \lim_{r\to 0} \frac{v_s(B(x,r))}{u(B(x,r))} & \text{if, for all } r > 0,\ u(B(x,r)) > 0 \\ +\infty & \text{else.} \end{cases}$$

Hence,

$$D_u v = D_u v_{ac} + D_u v_s$$

$D_u v_s$ is zero $u-$ almost everywhere and hence,

$$D_u v = D_u v_{ac},\ u- \text{ almost everywhere}$$

Therefore, for A, a borel set ,

$$v_{ac}(A) = \int_A D_u v(x) du(x)$$

Since $v = v_{ac} + v_s$,

$$\implies v(A) = v_{ac}(A) + v_s(A) = \int_A D_u v(x) du(x) + v_s(A)$$

Thus,

$$v(A) = \int_A D_u v\ du + v_s(A)$$

For the case when $v(\mathbb{R}^n)$ is not finite, let $\{K_l\}$ be an increasing sequence of compact sets covering $\mathbb{R}^n$, that is

- $\mathbb{R}^n = \bigcup_{l=1}^{\infty} K_l$
- $K_l \subset K_{l+1}$

By the assumption, since v is a radon measure, $v(K_l) < \infty$, for all l. Let $v_l := (v|K_l)$. Then, by the previous case of the finite radon measure, there are borel sets $\widetilde{B_l} \in \epsilon$ such that

$$v_l(\widetilde{B_l}) = \inf_{B\in\epsilon} v_l(B)$$

Define

$$B_l := \bigcap_{j=l}^{\infty} \widetilde{B_j}$$

$$\implies B_l \subset B_{l+1}\ ,\ B_l \text{ is borel.}\ ,\ B_l \subset \widetilde{B_l} \text{ and}$$

$$\mathbb{R}^n - B_l = \mathbb{R}^n - \left(\bigcap_{j=l}^{\infty} \widetilde{B_j}\right) \subset \bigcup_{j=l}^{\infty} \left(\mathbb{R}^n - \widetilde{B_j}\right)$$

$$\implies u\big(\mathbb{R}^n - B_l\big) \leq \sum_{j=l}^{\infty} u\big(\mathbb{R}^n - \widetilde{B_j}\big) = 0$$

$$\implies B_l \in \epsilon \text{ and } \inf_{B\in\epsilon} v_l(B) \leq v_l(B_l) \leq v_l(\widetilde{B_l}) = \inf_{B\in\epsilon} v_l(B)$$

$$\implies v_l(B_l) = v_l(\widetilde{B_l}) = \inf_{B\in\epsilon} v_l(B)$$

Let

$$B_0 := \bigcup_{l=1}^{\infty} B_l$$

$$v_{ac} := (v|B_0)$$

$$v_s := (v|\mathbb{R}^n - B_0)$$

Then, for any borel set A,

$$v(A) = v(A \cap B_0) + v\big(A \cap (\mathbb{R}^n - B_0)\big) = v_{ac}(A) + v_s(A)$$

Clearly, $v_s(B_0) = u(\mathbb{R}^n - B_0) = 0 \implies v_s \perp u$.
Claim : $v_{ac} \ll u$.
Proof of the claim : Let $u(A) = 0$, A is a borel set. Then

$$A \cap B_0 = \phi \implies v_{ac}(A) = v(A \cap B_0) = 0$$

Now, consider the case $A \subset B_0$.
Suppose that $v_{ac}(A) > 0$, then $v(A \cap B_0) > 0$. Since $A = \cup_{n=1}^{\infty}\big(A \cap K_n\big)$, we have $v\big(\cup_{m=1}^{\infty} \cup_{l=1}^{\infty} (A \cap K_m \cap B_l)\big) > 0$. Hence, there are n_0, l_0 such that $v(A \cap K_{n_0} \cap B_{l_0}) > 0$. Since, $K_m \subset K_{m+1}$, $B_l \subset B_{l+1}$,

$$\implies \forall m \geq n_0,\ l \geq l_0,\ v\big(A \cap K_m \cap B_l\big) > 0$$

Choose $p \geq \max\{n_0, l_0\}$, then $v\big(A \cap K_p \cap B_p\big) > 0$.

$$u\Bigg(\mathbb{R}^n - \big(B_p - (A \cap B_p)\big)\Bigg) = u\big(\mathbb{R}^n - B_p\big) + u(A \cap B_p) \leq u\big(\mathbb{R}^n - B_p\big) + u(A) = 0$$

$$\implies B_p - (A \cap B_p) \in \epsilon$$

$$\implies \inf_{B \in \epsilon} v_p(B) \leq v_p\Bigg(B_p - (A \cap B_p)\Bigg) = v_p(B_p) - v_p(A \cap B_p) = \inf_{B \in \epsilon} v_p(B) - v\big(A \cap K_p \cap B_p\big)$$

$$< \inf_{B \in \epsilon} v_p(B)$$

This is a clear contradiction. Hence, $v_{ac}(A) = 0$ and thus $v_{ac} \ll u$.
Since, $v_s(B_0) = 0$, $u\big(\mathbb{R}^n - B_0\big) = 0$, we have, as noted earlier, $D_u v_s(x) = 0$ for u a.e x. Hence, for u a.e x,

$$D_u v(x) = D_u v_{ac}(x) + D_u v_s(x) = D_u v_{ac}(x)$$

$$\implies v(A) = v_{ac}(A) + v_s(A) = \int_A \big(D_u v_{ac}\big)(x) du(x) + v_s(A)$$

$$= \int_A D_u v(x) du(x) + v_s(A)$$

□

Definition 2.17.

- Denote the average of f over the set E, with respect to u as

$$\fint_E f \ du := \frac{1}{u(E)} \int_E f \ du$$

  The definition is valid provided $0 < u(E) < \infty$ and the intergal on the $R.H.S$ is defined.

- $\mathcal{L}^1_{loc}(\mathbb{R}^n, u) := \{f : \mathbb{R}^n \to \mathbb{R} \ ; \ \int_K |f| du < \infty, \text{ for all compact sets, } K \subset \mathbb{R}^n \ \}$

Theorem 2.18. ***Lebesgue Besicovitch Differentiation theorem***
Let u be a radon measure on $\mathbb{R}^n$, $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n, u)$. Then,

$$\lim_{r \to 0} \fint_{B(x,r)} f \ du = f(x), \text{ for } u- \text{ a.e } x \in \mathbb{R}^n$$

Proof. For Borel $B \subset \mathbb{R}^n$, define

$$v^{\pm}(B) := \int_B f^{\pm} \ du$$

Then, by MCT, $v^{\pm}$ are $\sigma-$ finite measures on the $\sigma-$ algebra of borel sets. For an arbitrary set $A \subset \mathbb{R}^n$, define

$$v^{\pm}(A) \equiv \inf\{v^{\pm}(B) \ ; \ A \subset B, B \text{ is borel}\}$$

From the construction of measures from semi-algebras, it follows that $v^{\pm}$ are regular, borel measures. Now for any compact set K,

$$v^{\pm}(K) = \int_K f^{\pm} du(x) \leq \int_K |f| du < \infty$$

Thus $v^{\pm}$ are radon measrues as well.
Also, note that $v^{\pm} \ll u$. Thus, by the Radon-Nikodym theorem, for all $u-$ measurable set A,

$$\int_A f^+ \ du = v^+(A) = \int_A D_u v^+ \ du$$

$$\int_A f^- \ du = v^-(A) = \int_A D_u v^- \ du$$

Thus

$$D_u v^{\pm} = f^{\pm} \ ; \ u- \text{ almost everywhere}$$

Consequently,

$$\lim_{r \to 0} \fint_{B(x,r)} f \ du = \lim_{r \to 0} \fint_{B(x,r)} f^+ \ du - \lim_{r \to 0} \fint_{B(x,r)} f^- \ du$$

$$= \lim_{r \to 0} \frac{1}{u(B(x,r)} \left[v^+(B(x,r)) - v^-(B(x,r)) \right]$$

$$= D_u v^+(x) - D_u v^-(x) \ , \ u- \text{ almost everywhere}$$

$$= f^+(x) - f^-(x)$$

$$= f(x) \text{ for all } x, \ u- \text{ almost everywhere}$$

□

Remark 2.19. The same result holds for f to be in $\mathcal{L}^p_{loc}(\mathbb{R}^n, u)$, where

$$\mathcal{L}^p_{loc}(\mathbb{R}^n, u) := \{f : \mathbb{R}^n \to \mathbb{R} \ ; \ \int_K |f|^p du < \infty, \text{ for all compact sets, } K \subset \mathbb{R}^n \ \}$$

Corollary 2.20. *Let u be a radon measure on $\mathbb{R}^n$, $1 \leq p < \infty$ and $f \in \mathcal{L}^p_{loc}(\mathbb{R}^n, u)$ where*

$$\mathcal{L}^p_{loc}(\mathbb{R}^n, u) := \{f : \mathbb{R}^n \to \mathbb{R} \ ; \ \int_K |f|^p du < \infty, \textit{ for all compact sets, } K \subset \mathbb{R}^n \ \}$$

Then

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)|^p \; du(y) = 0 \text{ ; for } u- \text{ almost every } x$$

Definition 2.21. x is a lebesgue point with respect to u, if

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)|^p \; du(y) = 0$$

Proof of the corollary: Let $\{r_i\}_{i=1}^{\infty}$ *be countable dense set in* $\mathbb{R}$.
Claim 1 : $f(x) - r_i \in \mathcal{L}^p_{loc}(\mathbb{R}^n, u)$
Proof of the claim 1 : Fix a compact set $K \subset \mathbb{R}^n$*. Then, by the minikowski's inequality,*

$$\left(\int_K |f - r_i|^p \; du\right)^{\frac{1}{p}} \leq \left(\int_K |f|^p \; du\right)^{\frac{1}{p}} + \Big(|r_i|^p \; u(K)\Big)^{\frac{1}{p}} < \infty$$

This proves the claim 1.
By the Lebesgue-Besicovitch differentiation theorem 2.18,

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - r_i|^p \; du(y) = |f(x) - r_i|^p \text{ ; for } u- \text{ almost every } x.$$

This implies that there are $\{A_i \subset \mathbb{R}^n\}_{i=1}^{\infty}$ *such that, for* $A = \bigcup_{i=1}^{\infty} A_i$*, we have* $u(A) = 0$ *and satisfy the property:* $\forall x \notin A$, $\forall r_i$,

$$\Longrightarrow \lim_{r\to 0} \fint_{B(x,r)} |f(y) - r_i|^p \; du(y) = |f(x) - r_i|^p$$

Let $x \notin A$ *and by the dense property, choose* r_i *such that*

$$|f(x) - r_i|^p < \frac{\epsilon}{2^p}$$

Then, by the hölder's inequality,

$$\begin{aligned} &\limsup_{r\to 0} \fint_{B(x,r)} |f(y) - f(x)|^p \; du(y) \\ &\leq 2^{p-1} \left[\limsup_{r\to 0} \fint_{B(x,r)} |f(y) - r_i|^p \; du(y) + \fint_{B(x,r)} |f(x) - r_i|^p \; du(y)\right] \\ &= 2^{p-1} \Big[|f(x) - r_i|^p + |f(x) - r_i|^p\Big] \\ &< \epsilon \end{aligned}$$

ϵ *was arbitrarily greater than zero. This concludes the corollary.* $\square$

Corollary 2.22. *If $f \in \mathcal{L}^p_{loc}$ for some $1 \leq p < \infty$, then*

$$\lim_{B \downarrow \{x\}} \fint_B |f(y) - f(x)|^p \ dy = 0 \ for \ \mathcal{L}^n \ a.e \ x$$

Remark 2.23. Here $\lim\limits_{B \downarrow \{x\}}$ means the limit is taken over all closed balls containing x with $diam(B) \to 0$.
Note that the balls need not be centered at x in this kind of limit.

Proof. Let B be a ball with $diam(B) = d$ and $x \in B$ with x being the lebesgue point. Then, $B \subset B(x,d)$ and hence,

$$\frac{\mathcal{L}^n\big(B(x,d)\big)}{\mathcal{L}^n(B)} = \frac{\alpha(n)d^n}{\alpha(n)\left(\frac{d}{2}\right)^n} = 2^n$$

Hence,

$$\fint_B |f(y) - f(x)|^p dy \leq \frac{\mathcal{L}^n\big(B(x,d)\big)}{\mathcal{L}^n(B)} \fint_{B(x,d)} |f(y) - f(x)|^p dy \leq 2^n \fint_{B(x,d)} |f(y) - f(x)|^p dy$$

As x is a lebesgue point, $d \to 0 \implies RHS \to 0$ and this proves the corollary. ∎

Corollary 2.24. *Let $E \subset \mathbb{R}^n$ be $\mathcal{L}^n-$ measurable, then*

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 1 \ for \ \mathcal{L}^n- \ almost \ every \ x \in E$$

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 0 \ for \ \mathcal{L}^n- \ almost \ every \ x \in E^c$$

Proof. Let $f \equiv \mathcal{X}_E$. As E is $\mathcal{L}^n-$ measurable, f is locally integrable as $u(E \cap K) \leq u(K) < \infty$. By the above corollary,

$$\lim_{r \to 0} \fint_{B(x,r)} f(y) \ dy = f(x) \text{ for } \mathcal{L}^n- \text{ almost every x.}$$

As per the hypothesis,

$$L.H.S \equiv \lim_{r \to 0} \frac{\mathcal{L}^n(B(x,r) \cap E)}{\mathcal{L}^n(B(x,r))} = 0 \text{ or } 1 \equiv \text{R.H.S accordingly a.e.}$$

This proves the corollary. ∎

Definition 2.25. Points of density :
Let E be a $\mathcal{L}^n-$ measurable set of $\mathbb{R}^n$. $x \in \mathbb{R}^n$ is called a point of density for E if

$$\lim_{r\to 0} \frac{\mathcal{L}^n\big(B(x,r)\cap E\big)}{\mathcal{L}^n\big(B(x,r)\big)} = 1$$

Remark 2.26. From the previous corollary, for almost every $x \in E$, x is a point of density.

2.2.2 Riesz Representation Theorem.

Definition :

$$C_c(\mathbb{R}^n;\mathbb{R}^m) := \{f : \mathbb{R}^n \to \mathbb{R}^m \ ; \ f \text{ is continous and has compact support in } \mathbb{R}^n\}$$

Theorem 2.27. *Let $L : C_c(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}$ be a linear functional satisfying, for each compact set $K \subset \mathbb{R}^n$,*

$$\sup\{L(f) \ ; \ f \in C_c(\mathbb{R}^n;\mathbb{R}^m) \ ; |f| \le 1 \ , \ support(f) \subset K\} < \infty \qquad --(*)$$

Then there is a radon measure u on $\mathbb{R}^n$ and a $u-$ measurable function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ such that for all $f \in C_c(\mathbb{R}^n;\mathbb{R}^m)$,

- $|\sigma(x)| = 1$ *for $u-$ almost every x*
- $L(f) = \int_{\mathbb{R}^n} f\sigma \ du$

Definition 2.28. u is called the variational measure which is defined for each open set $V \subset \mathbb{R}^n$ as

$$u(V) := \sup\{L(f) \ ; \ f \in C_c(\mathbb{R}^n;\mathbb{R}^m) \ ; \ |f| \le 1 \ , \ support(f) \subset V\}$$

Proof. Firstly observe that $L(f) < 0 \implies L(-f) > 0$. Hence, for an open set, define u on open sets V as above and set and for an arbitrary set $A \subset \mathbb{R}^n$

$$u(A) \equiv \inf\{u(V) \ ; \ A \subset V; V \text{ is open}\}$$

Claim 1 : u is a measure (i.e, subadditivity is satisfied).
Proof of the claim 1 : Let V , $\{V_i\}_{i=1}^{\infty}$ be open sets in $\mathbb{R}^n$ with $V \subset \cup_{i=1}^{\infty} V_i$. Choose $g \in C_c(\mathbb{R}^n;\mathbb{R}^m)$ such that $|g| \le 1$ and $support(g) \subset V$.
By the compactness, $support(g) \subset \cup_{j=1}^{k} V_j$. By the partitions of unity, let $\{\xi_j\}_{j=1}^{k}$ be a finite sequence of smooth, non negative functions such that $support(\xi_j) \subset V_j$ and

$\sum_{j=1}^{k} \xi_j \equiv 1$ on $support(g)$. Then, clearly, $g \equiv \sum_{j=1}^{k} g\xi_j$ and hence,

$$|L(g)| = \left| \sum_{j=1}^{k} L(g\xi_j) \right| \leq \sum_{j=1}^{k} |L(g\xi_j)| \leq \sum_{j=1}^{\infty} u(V_j)$$

Taking supremum over g, we get

$$u(V) \leq \sum_{j=1}^{\infty} u(V_j)$$

Now let $\{A_j\}_{j=1}^{\infty}$ be arbitrary sets with $A \subset \cup_{j=1}^{\infty} A_j$. Fix $\epsilon > 0$. Choose open sets V_j such that $A_j \subset V_j$ and $u(A_j) + \frac{\epsilon}{2^j} \geq u(V_j)$. Then

$$u(A) \leq u\left(\bigcup_{j=1}^{\infty} V_j\right) \leq \sum_{j=1}^{\infty} u(V_j) \leq \sum_{j=1}^{\infty} u(A_j) + \epsilon$$

This proves the claim 1 as ϵ was arbitrarily positive.
Claim 2 : u is a radon measure.
Proof of the claim 2: Let U, V be two open sets with $dist(U, V) > 0$.
Let $g \in C_c\Big(U \cup V, \mathbb{R}^m\Big)$ and $|g| \leq 1$. Then $g_1 := g|_U$ and $g_2 := g|_V$ are in $C_c(U, \mathbb{R}^m)$ and $C_c(V, \mathbb{R}^m)$ respectively. Also, $g = g_1 + g_2$ with $|g_1| \leq 1$ and $|g_2| \leq 1$. Hence,

$$L(g) = L(g_1) + L(g_2) \leq u(U) + u(V)$$

$$\text{Taking supremum over } g \implies u\Big(U \cup V\Big) \leq u(U) + u(V)$$

Let $\epsilon > 0$. Choose g_1, g_2 with $g := g_1 + g_2$ and $u(U) \leq L(g_1) + \epsilon/2$ and $u(V) \leq L(g_2) + \epsilon/2$

$$\implies u\Big(U \cup V\Big) \geq L(g) = L(g_1) + L(g_2) \geq u(U) + u(V) - \epsilon$$

Letting $\epsilon \to 0$,we have

$$u(U \cup V) = u(U) + u(V)$$

The above is true for open sets. Now, for arbitrary 2 sets with strict positive distance between them, we have :
Let A, B be subsets of $\mathbb{R}^m$ and $dist(A, B) = r > 0$.
Let $\epsilon = \frac{r}{4}$ and $U_\epsilon := \{x \in \mathbb{R}^n \ ; \ d(x, A) < \epsilon\}$ and $V_\epsilon := \{x \in \mathbb{R}^n \ ; \ d(x, B) < \epsilon\}$. Then $d(U_\epsilon, V_\epsilon) \geq \frac{r}{2}$. Let $A \cup B \subset W_\epsilon$, where W_ϵ is an open set such that $u\big(A \cup B\big) \geq u\big(W_\epsilon\big) - \frac{\epsilon}{2}$. Declare $W_{1,\epsilon} := W_\epsilon \cap U_\epsilon$ and $W_{2,\epsilon} = W_\epsilon \cap V_\epsilon$, then $dist(W_{1,\epsilon}, W_{2,\epsilon}) \geq dist(U_\epsilon, V_\epsilon) \geq \frac{r}{2}$

$$\implies u\big(A \cup B\big) \geq u(W_\epsilon) - \frac{\epsilon}{2} \geq u\big(W_{1,\epsilon} \cup W_{2,\epsilon}\big) - \frac{\epsilon}{2}$$

$$\geq u(W_{1,\epsilon}) + u(W_{2,\epsilon}) - \frac{\epsilon}{2} \geq u(A) + u(B) - \frac{\epsilon}{2}$$

Letting $\epsilon \to 0$, we have

$$u(A) + u(B) \leq u(A \cup B) \leq u(A) + u(B)$$

$$\implies u(A \cup B) = u(A) + u(B)$$

Thus, by the caratheodory's criteria, we have u to be a borel measure. Also, by the definition of u, it is also borel regular, i.e, given $A \subset \mathbb{R}^n$, there is V_k, open such that $A \subset V_k$ and $u(V_k) \leq u(A) + \frac{1}{k}$, for all $k \in \mathbb{N}$. Thus $u(A) = u(\cap_{k=1}^{\infty} V_k)$. Now, $(*)$ condition of the hypothesis implies that $u(K) < \infty$, for all K, compact subsets of $\mathbb{R}^n$. This proves the claim 2.

Now, let

$$C_c^+(\mathbb{R}^n) := \{f \in C_c(\mathbb{R}^n, \mathbb{R}) \ ; \ f \geq 0\}$$

For $f \in C_c^+(\mathbb{R}^n)$, set

$$\lambda(f) := \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq f\}$$

Some observations include :

- If $f, g \in C_c^+(\mathbb{R}^n)$ and $f \leq g$, then $\lambda(f) \leq \lambda(g)$.
- For $f \in C_c^+(\mathbb{R}^n), \lambda(cf) = c\lambda(f)$, for all $c \geq 0$.

Claim 3 : For all $f, g \in C_c^+(\mathbb{R}^n)$, we have $\lambda(f+g) = \lambda(f) + \lambda(g)$.
Proof of the claim 3 : If $h_1, h_2 \in C_c(\mathbb{R}^n, \mathbb{R}^m)$ with $|h_1| \leq f$ and $|h_2| \leq g$,
then $|h_1 + h_2| \leq f + g$. We can, without loss of generality, assume that $L(h_1), L(h_2) \geq 0$. Thus

$$|L(h_1)| + |L(h_2)| = L(h_1) + L(h_2) = L(h_1 + h_2) = |L(h_1 + h_2)| \leq \lambda(f+g)$$

Taking supremum over h_1 and h_2 in $C_c(\mathbb{R}^n, \mathbb{R}^m)$, we get

$$\lambda(f) + \lambda(g) \leq \lambda(f+g)$$

Now fix $h \in C_c(\mathbb{R}^n, \mathbb{R}^m)$ with the criteria that $|h| \leq f + g$. Set

$$h_1 \equiv \begin{Bmatrix} \frac{f.h}{f+g} & \text{if } f+g > 0 \\ 0 & \text{if } f+g = 0 \end{Bmatrix}$$

$$h_2 \equiv \begin{Bmatrix} \frac{g.h}{f+g} & \text{if } f+g>0 \\ 0 & \text{if } f+g=0 \end{Bmatrix}$$

Then, clearly, $h_1, h_2 \in C_c(\mathbb{R}^n, \mathbb{R}^m)$. Also $h = h_1 + h_2$ and $|h_1| \leq f$ and $|h_2| \leq g$.Thus

$$|L(h)| \leq |L(h_1)| + |L(h_2)| \leq \lambda(f) + \lambda(g)$$

As a consequence,

$$\lambda(f+g) \leq \lambda(f) + \lambda(g)$$

This proves the claim 3.

Claim 4 : $\lambda(f) = \int_{\mathbb{R}^n} f \; du$; for all $f \in C_c^+(\mathbb{R}^n)$.

Proof of the claim 4 : First observe that $f^{-1}(\{t\})$ is a closed set in $\mathbb{R}^n$ and $f^{-1}(\{t\}) \cap f^{-1}(\{s\}) = \phi$ if $t \neq s$. Let $0 < s < t$ and $K \subset \mathbb{R}^n$ be a compact set. For every $l > 0$, define

$$S_l(K) := \{\theta \in [s,t] \; ; \; u\big(f^{-1}(\theta) \cap K\big) \geq \frac{1}{l}\}$$

$$S(K) := \{\theta \in [s,t] \; ; \; u\big(f^{-1}(\theta) \cap K\big) > 0\}$$

Clearly,

$$S(K) = \bigcup_{l=1}^{\infty} S_l(K)$$

Since u is a radon measure, $u\big(f^{-1}(\theta) \cap K\big) \leq u(K) < \infty$

$$\implies \frac{card\big(S_l(K)\big)}{l} \leq \sum_{\theta \in S_l(K)} u\big(f^{-1}(\theta) \cap K\big) = u\Bigg(\bigcup_{\theta \in S_l(K)} f^{-1}(\theta) \cap K\Bigg) \leq u(K)$$

$$\implies card\big(S_l(K)\big) \leq l \; u(K) < \infty$$

Hence, for every K compact, $S(K)$ is countable. Let $S := \{\theta \in [s,t] \; ; \; u\big(f^{-1}(\theta)\big) > 0\}$, then $S = \bigcup_{l=1}^{\infty} S(K_l)$, where $K_l \subset K_{l+1}$, $\mathbb{R}^n = \bigcup_{l=1}^{\infty} K_l$. Hence S is countable and $S \subset [s,t]$. Therefore, for almost every $\theta \in [s,t]$, $u\big(f^{-1}(\theta)\big) = 0$. Let $\epsilon > 0$. Choose $0 = t_0 < t_1 < t_2 < \ldots t_N$ such that $t_N = 2 \; ||f||_{L^\infty}$, $0 < t_i - t_{i-1} < \epsilon$, $u\left(f^{-1}\{t_i\}\right) = 0$ for $j = 1, 2 \ldots, N$. Declare

$$U_j := f^{-1}((t_{j-1}, t_j))$$

U_j's are clearly open and since support of f is compact, $u(U_j) < \infty$. By the approximation by the compact sets for radon measure, there is $K_j \subset U_j$ such that $u(U_j - K_j) < \frac{\epsilon}{N}$, for all $j = 1, 2, \ldots, N$. Also, there are functions $g_j \in C_c(\mathbb{R}^n, \mathbb{R}^m)$ with $|g_j| \leq 1$, $support(g_j) \subset U_j$ and $|L(g_j)| \geq u(U_j) - \frac{\epsilon}{N}$. Also, observe that there are functions $h_j \in C_c^+(\mathbb{R}^n)$ such that $support(h_j) \subset U_j$ and $0 \leq h_j \leq 1$, and $h_j \equiv 1$ on

$K_j \cup support(g_j)$. Then,

$$\lambda(h_j) \geq |L(g_j)| \geq u(U_j) - \frac{\epsilon}{N}$$

and

$$\lambda(h_j) = \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq h_j\}$$

$$\leq \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq 1 \ , \ support(g) \subset U_j\}$$

$$= u(U_j)$$

whence

$$u(U_j) - \frac{\epsilon}{N} \leq \lambda(h_j) \leq u(U_j)$$

Since $\{U_j\}$ are disjoint and $support(h_j) \subset U_j \implies f\big(1 - \sum_{j=1}^{N} h_j\big) \geq 0$, declare

$$A := \{x \ ; \ f(x)\left(1 - \sum_{i=1}^{N} h_i(x)\right) > 0\}$$

Here, A is clearly open. Now, we compute

$$\lambda\left(f - f\ \sum_{j=1}^{N} h_j\right) = \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq \big(f - f\sum_{j=1}^{N} h_j\big)\mathcal{X}_A\}$$

$$\leq \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq ||f||_{L^\infty}\ \mathcal{X}_A\}$$

$$= ||f||_{L^\infty} \sup\{|L(g)| \ ; \ g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \ , \ |g| \leq \mathcal{X}_A\}$$

$$= ||f||_{L^\infty}\ u(A)$$

$$= ||f||_{L^\infty}\ u\left(\bigcup_{j=1}^{N}(U_j - \{h_j = 1\})\right)$$

$$\leq ||f||_{L^\infty} \sum_{j=1}^{N} u(U_j - K_j)$$

$$\leq \epsilon ||f||_{L^\infty}$$

Hence,

$$\lambda(f) = \lambda\left(f - f\sum_{j=1}^{N} h_j\right) + \lambda\left(f\sum_{j=1}^{N} h_j\right)$$

$$\leq \epsilon ||f||_{L^\infty} + \sum_{j=1}^{N} \lambda(f.h_j)$$

$$\leq \epsilon ||f||_{L^\infty} + \sum_{j=1}^{N} t_j.u(U_j)$$

and

$$\lambda(f) \geq \sum_{j=1}^{N} \lambda(f.h_j)$$

$$\geq \sum_{j=1}^{N} t_{j-1} \left(u(U_j) - \frac{\epsilon}{N} \right)$$

$$\geq \sum_{j=1}^{N} t_{j-1} u(U_j) - t_N \epsilon$$

Since $u\big(f^{-1}(\{t_j\})\big) = 0$,

$$\sum_{j=1}^{N} t_{j-1} u(U_j) \leq \int_{\mathbb{R}^n} f \; du \leq \sum_{j=1}^{N} t_j u(U_j)$$

we have

$$|\lambda(f) - \int_{\mathbb{R}^n} f \; du| \leq \sum_{j=1}^{N} (t_j - t_{j-1}) u(U_j) + \epsilon ||f||_{L^\infty} + \epsilon t_N$$

$$\leq \epsilon \; u(support(f)) + 3 \; \epsilon ||f||_{L^\infty}$$

This proves the claim 4.

Claim 5 : There is a $u-$ measurable function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ such that $L(f) = \int_{\mathbb{R}^n} f.\sigma \; du$ for all $f \in C_c(\mathbb{R}^n, \mathbb{R}^m)$.
Proof of the claim 5 : Fix $a \equiv (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$ such that $|a| = 1$. Define

$$\lambda_a(f) \equiv L(fa)$$

Note that $(f \; a)(x) \equiv f(x).(a_1, a_2, \ldots, a_m) \equiv (a_1 f(x), a_2, f(x), \ldots, f(x) a_m) \in \mathbb{R}^m$ and is compactly supported, for $f \in C_c(\mathbb{R}^n)$. Then λ_a is linear and

$$|\lambda_a(f)| = |L(f \; a)| \leq \sup\{|L(g)| \; ; \; g \in C_c(\mathbb{R}^n, \mathbb{R}^m) \; , \; |g| \leq |f|\}$$

$$= \lambda(|f|) = \int_{\mathbb{R}^n} |f| \; du$$

Thus,by the Hanhn-Banach theorem, we can extend λ_a to a bounded linear functional on $L^1(\mathbb{R}^n, u)$. Hence, there is $\sigma_a \in L^\infty(u)$ such that

$$\lambda_a(f) = \int_{\mathbb{R}^n} f \sigma_a \; du \; ; \text{for } f \in C_c(\mathbb{R}^n)$$

Let $e_1, e_2, \ldots e_m$ be the standard basis of $\mathbb{R}^m$. Define

$$\sigma \equiv \sum_{j=1}^{\infty} \sigma_{e_j} e_j$$

Then if $f \in C_c(\mathbb{R}^n, \mathbb{R}^m)$, we have

$$L(f) = \sum_{j=1}^{m} L((f.e_j)e_j)$$

$$= \sum_{j=1}^{m} \int (f.e_j)\sigma_{e_j} \; du$$

$$= \int f.\sigma \; du$$

This proves the claim 5.

Claim 6 : $|\sigma| = 1$ $u-$ a.e.

Proof of the claim 6 : Let $U \subset \mathbb{R}^n$ be open and $u(U) < \infty$. Then, by definition,

$$u(U) = \sup\{\int f.\sigma \; du \; ; \; f \in C_c(\mathbb{R}^n, \mathbb{R}^m) \; , \; |f| \leq 1 \; , \; support(f) \subset U\} \qquad (**)$$

Take $g_k \in C_c(\mathbb{R}^n, \mathbb{R}^m)$ such that $|g_k| \leq 1$ and $support(g_k) \subset U$ and $g_k.\sigma \to |\sigma|$; $u-$ almost everywhere. Then, by the $(**)$,

$$\int_U |\sigma| \; du = \lim_{k \to \infty} \int g_k.\sigma \; du \leq u(U)$$

Also, if $f \in C_c(\mathbb{R}^n, \mathbb{R}^m)$ with $|f| \leq 1$ and $support(f) \subset U$, then

$$\int f.\sigma \; du \leq \int_U |\sigma| \; du$$

By (**),

$$u(U) \leq \int_U |\sigma| \; du$$

Thus

$$u(U) = \int_U |\sigma| \; du$$

for all open sets $U \subset \mathbb{R}^n$ and hence,

$$|\sigma| = 1 \; ; \; u- \text{ almost everywhere}$$

$\square$

Corollary 2.29. *Let $L : C_c(\mathbb{R}^n) \to \mathbb{R}$ be linear and non-negative, so that*

$$L(f) \geq 0 \text{ for all } f \in C_c^\infty(\mathbb{R}^n), f \geq 0$$

Then there is a radon measure u on $\mathbb{R}^n$ such that

$$L(f) = \int_{\mathbb{R}^n} f \; du \text{ for all } f \in C_c^\infty(\mathbb{R}^n)$$

Proof. Choose any compact set K in $\mathbb{R}^n$ and select smooth function ξ such that ξ has compact support, $\xi \equiv 1$ on K and $0 \leq \xi \leq 1$. Then, for any $f \in C_c^\infty(\mathbb{R}^n)$ with $support(f) \subset K$, set $g \equiv ||f||_{L^\infty}\xi - f \geq 0$. The hypothesis of the corollary implies

$$0 \leq L(g) = ||f||_{L^\infty} L(\xi) - L(f)$$

and so

$$L(f) \leq C||f||_{L^\infty}$$

where $C \equiv L(\xi)$. Thus L extends to a linear mapping from $C_c(\mathbb{R}^n)$ to $\mathbb{R}$ such that it satisfies the condition of the Reisz representation theorem. Hence, there is u, σ as before such that

$$L(f) = \int_{\mathbb{R}^n} f.\sigma \; du \;\; f \in C_c^\infty(\mathbb{R}^n)$$

with $\sigma = \pm 1$ $u-$ a.e. The hypothesis of the corollary forces $\sigma = 1$, $u-$ a.e. □

Remark 2.30. There is a generalised version of the Riesz representation theorem in Real and Complex analysis, W.Rudin which says the following :

Let X be a locally compact Hausdorff space, λ be a positive linear functional on $C_c(X)$, i.e if $f \geq g$, then $\lambda(f) \geq \lambda(g)$. Then, there is a $\sigma-$ algebra $\mathcal{M}$ which contains all the borel sets in X and there is a unique positive measure u on $\mathcal{M}$ which represents λ such that

- $\lambda(f) = \int_X f \; du$ for all $f \in C_c(X) := \{f : X \to \mathbb{R} \; ; \; support(f) \text{ is compact in } X\}$
- $u(K) < \infty$ for all K compact subset of X
- For all $E \in \mathcal{M}$,

$$u(E) \equiv \inf\{u(V) \; ; \; E \subset V \; , \; V \text{ is open.}\}$$

- For all E open and in $\mathcal{M}$, with $u(E) < \infty$,

$$u(E) := \sup\{u(K) \; ; \; K \subset E \; , \; K \text{ is compact.}\}$$

- If $E \in \mathcal{M}$, $A \subset E$, $u(E) = 0 \implies A \in \mathcal{M}$

□

2.3 Hausdorff Measures

.

2.3.1 Definition and some elementary properties.

Definition 2.31. Let $A \subset \mathbb{R}^n$, $0 \leq s < \infty$, $0 \leq \delta < \infty$. Define

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

where

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

$$\Gamma(s) := \int_0^\infty e^{-x} x^{s-1} dx \; ; \; 0 < s < \infty$$

From the definition, $\alpha(0) = 1$ and $\alpha(1) = \frac{\pi^{\frac{1}{2}}}{\Gamma(\frac{1}{2})}$. Now,

$$\Gamma(\frac{1}{2}) = \int_0^\infty e^{-t} t^{\frac{-1}{2}} dt = 2\int_0^\infty e^{-x^2} dx$$

$$\implies \Gamma(\frac{1}{2})^2 = 4\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dx \; dy = 4\int_0^{\frac{\pi}{2}} \left(\int_0^\infty r e^{-r^2} dr\right) d\theta = \pi^{\frac{1}{2}}$$

$$\implies \alpha(1) = 2$$

Definition 2.32. For A and s as above, define

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

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

Remark 2.33. Here ‘supremum’ is the ‘limit’ as $\mathcal{H}_\delta^s$ increases as $\delta \to 0$, by the infimum property.

Theorem 2.34. *$\mathcal{H}^s$ is a Borel - regular measure. ($0 \leq s < \infty$)*

Remark 2.35. $\mathcal{H}^s$ need not be a radon measure as for $\mathcal{H}^0$ is counting measure, $\mathcal{H}^0([0,1])$ is not finite.

Proof. Claim 1 : $\mathcal{H}^s_\delta$ is a measure.
Proof of the claim 1 : Let $\{A_k\}_{k=1}^\infty \subset \mathbb{R}^n$ and suppose that $A_k \subset \bigcup_{j=1}^\infty C_j^k$, for all k with $diam(C_j^k) \le \delta$, then $\{C_j^k\}_{j,k=1}^\infty$ becomes a cover for $\bigcup_{k=1}^\infty A_k$.
Thus

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

Taking the infimum, we get

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

And clearly, if $A \subset B$, then,

$$\mathcal{H}^s_\delta(A) \le \mathcal{H}^s_\delta(B) \text{ and } \mathcal{H}^s_\delta(\phi) = 0$$

This proves the claim 1.

Claim 2 : $\mathcal{H}^s$ is a measure.
Proof of the claim 2 : Let $\{A_k\}_{k=1}^\infty \subset \mathbb{R}^n$. Then, by the claim 1,

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

The second part of the inequality is due to the supremum property. The above is true for all $\delta > 0$ and hence, taking $\delta \to 0$, we get

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

Clearly $\mathcal{H}^s(\phi) = 0$ as $\mathcal{H}^s_\delta(\phi) = 0$. Also, for $A \subset B$, $\mathcal{H}^s_\delta(A) \le \mathcal{H}^s_\delta(B) \le \mathcal{H}^s(B)$. hence, by letting $\delta \to 0$, we have $\mathcal{H}^s(A) \le \mathcal{H}^s(B)$.
This proves the claim 2.

Claim 3 : $\mathcal{H}^s$ is a Borel measure.
Proof of the claim 3 : (idea is to use the caratheodory criteria).
Choose $A, B \subset \mathbb{R}^n$ with $dist(A,B) > 0$. Let $0 < \delta < \frac{dist(A,B)}{4}$. Suppose $A \cup B \subset \bigcup_{k=1}^\infty C_k$ and $diam(C_k) \le \delta$, then , set

$$\widetilde{A} := \{C_j \ ; \ C_j \cap A \neq \phi\}$$

$$\widetilde{B} := \{C_j \ ; C_j \cap B \neq \phi\}$$

Clearly, if $C_j \in \widetilde{A}$ and $C_k \in \widetilde{B}$, then, $\big(C_j \cap A\big) \cap \big(C_k \cap B\big) = \phi$,
$diam(C_j \cap A) \leq diam(C_k) \leq \delta$, $diam(C_k \cap B) \leq diam(C_k) \leq \delta$, by the construction of δ.
Hence,

$$\sum_{j=1}^{\infty} \alpha(s).\left(\frac{diam(C_j)}{2}\right)^s \geq \sum_{C_j \in \widetilde{A}} \alpha(s).\left(\frac{diam(C_j \cap A)}{2}\right)^s + \sum_{C_j \in \widetilde{B}} \alpha(s).\left(\frac{diam(C_j \cap B)}{2}\right)^s$$

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

This is true for all covers $\{C_j\}$ of $A \cup B$ and thus

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

provided $0 < \delta < \frac{dist(A,B)}{4}$.
Let $\delta \to 0$, then,

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

By the sub-additivity, we have

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

provided $dist(A, B) > 0$.
By the caratheodory criteria, $\mathcal{H}^s$ is borel.
This proves the claim 3.

Claim 4 : $\mathcal{H}^s$ is Borel-regular.
Proof of the claim 4 : Observe that $diam(C) = diam(\overline{C})$, for all subsets C of $\mathbb{R}^n$. Hence

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

Let $A \subset \mathbb{R}^n$.
Case 1 : $\mathcal{H}^s(A) = \infty$:
As $A \subset \mathbb{R}^n$, $\mathcal{H}^s(\mathbb{R}^n) = \infty$.
Case 2 : $\mathcal{H}^s(A) < \infty$
$\implies \mathcal{H}^s_\delta(A) < \infty$, for all $\delta > 0$. For $k \geq 1$, choose $\{C^k_j\}_{j=1}^{\infty}$ such that

- $diam(C^k_j) \leq \frac{1}{k}$
- $A \subset \bigcup_{j=1}^{\infty} C^k_j$
- By the infimum property, $\sum_{j=1}^{\infty} \alpha(s).\left(\frac{diam(C^k_j)}{2}\right)^s \leq \mathcal{H}^s_{1/k}(A) + \frac{1}{k}$

Let $A_k := \cup_{j=1}^{\infty} C_j^k$ and $B := \cap_{k=1}^{\infty} A_k$

Then $A \subset A_k$ for all $k \implies A \subset B \implies \mathcal{H}^s(A) \leq \mathcal{H}^s(B)$.

Also, for all k, $B \subset A_k = \bigcup_{j=1}^{\infty} C_j^k$, hence

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

Sending $k \to \infty$, we get

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

And hence, for all A, there is B Borel subset of $\mathbb{R}^n$,

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

This proves the claim 4 and thus the theorem.

Remark: For $\mathcal{G}_\delta$ to be the collection of sets which are $\mathcal{H}^s_\delta$ measurable, note that in general, for $\delta > 0$, $\mathcal{G}^s_\delta$ need not contain the Borel sigma algebra.

Counter-Example : Let $X = \mathbb{R}$, $d(x,y) = |x-y|$ and $0 < s < 1$. Let $\mathcal{H}^s_\delta$ and $\mathcal{H}^s \equiv \lim_{\delta\to 0} \mathcal{H}^s_\delta$ be the corresponding hausdorff measures as defined earlier.

$$\text{As } \left(\sum_{i=1}^{\infty} \alpha_i\right)^s \leq \sum_{i=1}^{\infty} \alpha_i^s \text{ , if } 0 < s \leq 1$$

$\implies$ for any $[a,b]$ with $b - a \leq (\delta \equiv 1)$,

$$\mathcal{H}^s_1\big([a,b]\big) = \mathcal{H}^s_1\big((a,b]\big) = \mathcal{H}^s_1\big([a,b)\big) = \mathcal{H}^s_1\big((a,b)\big) = (b-a)^s$$

Hence,

$$\mathcal{H}^s_1\big([0,\frac{1}{2}]\big) = \frac{1}{2^s}$$

$$\mathcal{H}^s_1\big((\frac{1}{2},1]\big) = \frac{1}{2^s}$$

But,

$$\mathcal{H}^s_1\big([0,1]\big) \neq \mathcal{H}^s_1\big([0,\frac{1}{2}]\big) + \mathcal{H}^s_1\big((\frac{1}{2},1]\big)$$

Hence, the intervals are not measurable with respect to the measure $\mathcal{H}^s_1$. $\square$

Theorem 2.36. *Elementary properties of Hausdorff measures.*

- $\mathcal{H}^0$ *is counting measure.*
- $\mathcal{H}^1 = \mathcal{L}^1$ *on* $\mathbb{R}$
- $\mathcal{H}^s \equiv 0$ *on* $\mathbb{R}^n$ *, for all* $s > n$
- $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$*, for all* $\lambda > 0, A \subset \mathbb{R}^n$
- $\mathcal{H}^s(L(A)) = \mathcal{H}^s(A)$*, for all isometries* $L : \mathbb{R}^n \to \mathbb{R}^n$ *, that is, distance preserving maps from* $\mathbb{R}^n$ *to itself, with* $A \subset \mathbb{R}^n$

Proof. • Observe that $\alpha(s = 0) = \frac{1}{\Gamma(1)} \equiv 1$.
Thus $\mathcal{H}^0(\{a\}) = 1$ and this proves that $\mathcal{H}^0$ is a counting measure.

- Let $A \subset \mathbb{R}^n$ and $\delta > 0$. Note that $\Gamma\left(\frac{3}{2}\right) = \sqrt{\frac{\pi}{4}}$. Hence $\alpha(1) = 2$. Also, if $C \subset \mathbb{R}$, then there is an interval I with $C \subset I$ such that $diam(C) = diam(I)$. Thus

$$\mathcal{L}^1(A) = \inf\{\sum_j diam(I_j) \ ; \ A \subset \bigcup_{j=1}^{\infty} I_j \ ; \ I_j \text{ are intervals.}\}$$

$$= \inf\{\sum_j diam(C_j) \ ; \ A \subset \bigcup_{j=1}^{\infty} C_j\}$$

$$\leq \inf\{\sum_j diam(C_j) \ ; \ A \subset \bigcup_{j=1}^{\infty} C_j \ ; \ diam(C_j) \leq \delta\}$$

$$= \inf\{\sum_j \left(\frac{diam(C_j)}{2}\right) .\alpha(1) \ ; \ A \subset \bigcup_{j=1}^{\infty} C_j \ ; \ diam(C_j) \leq \delta\} = \mathcal{H}^1_\delta(A)$$

Conversely, set, for $k \in \mathbb{Z}$

$$I_k := [k\delta, (1+k)\delta]$$

Then $diam(C_j \cap I_k) \leq \delta$ and

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

Hence,

$$\mathcal{L}^1(A) = \inf\{\sum_{j=1}^{\infty} diam(C_j) \ ; \ A \subset \bigcup_{j=1}^{\infty} C_j\}$$

$$\geq \inf\{\sum_{j=1}^{\infty} \sum_{k=-\infty}^{+\infty} diam(C_j \cap I_k) \ ; \ A \subset \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} (C_j \cap I_k)\} \geq \mathcal{H}^1_\delta(A)$$

Thus

$$\mathcal{L}^1(A) = \mathcal{H}^1_\delta(A) \text{ ; for all } \delta > 0$$

$$\implies \mathcal{L}^1 \equiv \mathcal{H}^1 \text{ on } \mathbb{R}^n$$

- Fix $m \geq 1$, an integer.
  The unit cube in $\mathbb{R}^n$ which is $Q \equiv [0,1]^n := [0,1] \times [0,1] \times \ldots \times [0,1]_{\text{n times}}$ can be decomposed into cubes of side $\frac{1}{m}$ and diameter $\frac{n^{1/2}}{m}$. Therefore,

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

  Now, if $s > n$, and letting $m \to \infty$ tells that $\mathcal{H}^s \equiv 0$ on Q , for all $s > n$ on unit cube Q. Now $\mathbb{R}^m \subset \bigcup Q_i$, where Q_i's are unit cubes spanning the space. Thus

$$\mathcal{H}^s(\mathbb{R}^n) \leq \sum \mathcal{H}^s(Q_i) = 0$$

  concludes that $\mathcal{H}^s(\mathbb{R}^n) = 0$.
- For L to be an isometry,

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

$$= \inf\{\sum_{j=1}^{\infty} \alpha(s). \left(\frac{diam(C_j)}{2}\right)^s \;;\; L(A) \subset \bigcup_{j=1}^{\infty} L(C_j) \;,\; diam(L(C_j)) \leq \delta\}$$

  Let $D_j = L(C_j)$, then,

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

  Sending $\delta \to 0$, we get

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

  This also implies $\mathcal{H}^s$ is translational invariant.

$\square$

Lemma 2.37. ***A convinient way to verify that $\mathcal{H}^s$ vanishes on a set:***
Suppose $A \subset \mathbb{R}^n$ and $\mathcal{H}^s_\delta(A) = 0$ for some $0 < \delta \leq \infty$, then $\mathcal{H}^s(A) = 0$

Proof. Case 1 : $s \equiv 0$. Given that $\mathcal{H}^0_\delta(A) = 0$ for some $0 < \delta \leq \infty$,
Claim : $A = \phi$.
Proof of the claim : Suppose not, then, there is some $x \in A$. Consider C_j of diameter δ

(say a ball of radius $\frac{\delta}{2}$ at x). Note that $\alpha(s=0)=1$ and

$$\left(\frac{diam(C_j)}{2}\right)^{s=0} \equiv 1$$

and hence, $\mathcal{H}^s_\delta(A) \geq 1$ for $s=0 \implies 0 \geq 1$. This is a contradiction and thus proves the claim and hence the case 1.
Case 2 : $s > 0$ and $\mathcal{H}^s_\delta(A) = 0$ for some $0 < \delta \leq \infty$.
Fix $\epsilon > 0$. Then there is $\{C_j\}_{j=1}^\infty$ such that $diam(C_j) < \delta$. Note that the choice of C_j depends on ϵ as

$$\sum_j \alpha(s).\left(\frac{diam(C_j)}{2}\right)^s < \epsilon$$

$$\implies diam(C_j) \leq \left(\frac{\epsilon}{\alpha(s)}\right)^{\frac{1}{s}} \times 2$$

$$\implies diam(C_j) \leq \min\left\{\delta\ ,\ \left(\frac{\epsilon}{\alpha(s)}\right)^{\frac{1}{s}} \times 2\right\} \equiv \delta(\epsilon)$$

Hence,

$$\mathcal{H}^s_{\delta(\epsilon)} \leq \epsilon$$

Sending $\epsilon \to 0$, we have

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

□

Lemma 2.38. *Let $S \subset \mathbb{R}^n$ and $0 \leq s < t < \infty$.*

- *If $\mathcal{H}^s(A) < \infty$, then $\mathcal{H}^t(A) = 0$*
- *If $\mathcal{H}^t(A) > 0$, then $\mathcal{H}^s(A) = \infty$*

Proof. Note that the above 2 points are contrapositive to one another and hence, it suffices to prove just one of them, say the first one.
Let $\mathcal{H}^s(A) < \infty$ and $0 < \delta < 1$. It is needed to be shown that $\mathcal{H}^t(A) = 0$.

$$\implies \exists \{C_j\}_{j=1}^\infty \text{ such that } diam(C_j) \leq \delta\ ;\ A \subset \bigcup_{j=1}^\infty C_j$$

And

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

Then

$$\mathcal{H}^t_\delta(A) \leq \sum_{j=1}^\infty \alpha(t).\left(\frac{diam(C_j)}{2}\right)^t \ ;\text{ as } s \leq t$$

$$= \frac{\alpha(t)}{\alpha(s)} 2^{s-t} \sum_{j=1}^{\infty} \alpha(s). \left(\frac{diam(C_j)}{2} \right)^s diam(C_j)^{t-s}$$

$$\leq \frac{\alpha(t)}{\alpha(s)} \left(\mathcal{H}^s(A) + 1 \right) .\delta^{t-s}$$

As $\delta \to 0$, we have

$$\mathcal{H}^t_\delta(A) \to \mathcal{H}^t(A) = 0$$

Note that $\alpha(s) \neq 0$, for all s. □

Definition 2.39. The hausdorff dimension of a set $A \subset \mathbb{R}^n$ is defined to be

$$\mathcal{H}_{dim}(A) \equiv \inf\{0 \leq s < \infty \ ; \ \mathcal{H}^s(A) = 0\}$$

Remark 2.40.

- For $A \subset \mathbb{R}^n$, $\mathcal{H}_{dim}(A) \leq n$; by the elementary property (3) mentioned in theorem 2.36.
- Let $\mathcal{H}_{dim}(A) = s$, then

$$\mathcal{H}^t(A) = \left\{ \begin{array}{cc} 0 & \text{for all } t > s \\ \text{a real number} \equiv s & \text{for } t = s \\ +\infty & \text{for all } t < s \end{array} \right\}$$

- $\mathcal{H}_{dim}$ need not be an integer, in general.

2.3.2 Examples and calculation of Hausdorff measures for some sets.

- **Hausdorff dimension of the Cantor set.**
  The hausdorff dimension of the cantor set is $\frac{log2}{log3}$.

  Proof. Declare
  $$s = \frac{log2}{log3}$$

  The cantor set , as defined as the removal of the middle 1/3rd length spaces at each iteration, it can be denoted as $C := \cap_{k \in \mathbb{N}} C_k$, where each C_k is the finite union of the 2^k intervals of length $\frac{1}{3^k}$. Now, given $\delta > 0$, choose $K > 0$ such that $\frac{1}{3^K} < \delta$. Observe that C_K covers C and consists of 2^K intervals of length $\frac{1}{3^K} < \delta$. Thus

  $$\mathcal{H}^s_\delta(C) \leq \alpha \left(\frac{log2}{log3} \right) \left(\frac{2}{3^s} \right)^K$$

Note that s satisfies $3^s = 2$,

$$\implies \mathcal{H}^s(C) \leq \alpha\left(\frac{log2}{log3}\right)$$

Definition 2.41. A function f on a subset E of $\mathbb{R}^n$ satisfies a lipshitz/Hölder condition with exponent γ if there is a constant $M > 0$, such that for all $x, y \in E$,

$$|f(x) - f(y)| \leq M|x-y|^\gamma$$

Lemma 2.42. *Suppose that f is a function defined on a compact set E and satisfies the Hölder condition with the exponent γ, then, for some $M' > 0$,*

- *For $\beta = \frac{s}{\gamma}$, $\mathcal{H}^\beta\big(f(E)\big) \leq M'\mathcal{H}^s(E)$*
- *$dim\big(f(E)\big) \leq \frac{1}{\gamma}dim(E)$*

Proof. Suppose that $\{F_k\}_k$ is a countable family of sets that cover E, then $\{f(E \cap F_k)\}$ covers $f(E)$ and $diam\big(f(E \cap F_k)\big) \leq M\big(diam(F_k)\big)^\gamma$, where M is defined as in the definition 2.41. Thus,

$$\sum_{k=1}^{\infty}\bigg(diam\big(f(E \cap F_k)\big)\bigg)^{\frac{s}{\gamma}} \leq M^{\frac{s}{\gamma}}\sum_{i=1}^{\infty}\big(diam(F_i)\big)^s$$

Now, for $diam(F_k)$ less than δ and as $\delta \to 0$, we have $diam(f(E \cap F_k)) \to 0$ and hence, for $\beta \equiv \frac{s}{\gamma}$,

$$\mathcal{H}^\beta\big(f(E)\big) \leq M'\mathcal{H}^s(E)$$

where

$$M' = \frac{M^\beta 2^{s\left(1-\frac{1}{\gamma}\right)}\alpha(\beta)}{\alpha(s)}$$

This proves the first part of the lemma. The second part of the lemma follows from that fact that first part implies $\mathcal{H}^s(E) = 0 \implies \mathcal{H}^\beta(E) = 0$. $\square$

Remark 2.43. Construction of the cantor function :

As defined earlier, $C \equiv \cap_{k=1}^\infty C_k$, with C_k to be the union of disjoint intervals of length 2^k.

Iteration :

$C_1 \equiv [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$: Define f_1 continuous, on $[0,1]$ as

$$\left\{\begin{array}{cc} 0 & x = 0 \\ \frac{1}{2} & \frac{1}{3} \leq x \leq \frac{2}{3} \\ 1 & x = 1 \\ \text{A straight line joining the end points} & \text{elsewhere} \end{array}\right\}$$

$C_2 \equiv [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$: Define f_2 continuous, on $[0,1]$ as

$$\left\{\begin{array}{cc} 0 & x = 0 \\ \frac{1}{4} & \frac{1}{9} \leq x \leq \frac{2}{9} \\ \frac{1}{2} & \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{3}{4} & \frac{7}{9} \leq x \leq \frac{8}{9} \\ 1 & x = 1 \\ \text{A straight line joining the end points} & \text{elsewhere} \end{array}\right\}$$

And so on for each $i \in \mathbb{N}$.

Each f_i takes the value atmost 1 and f_i's are increasing. It also has the property that $|f_{n+1}(x) - f_n(x)| \leq \frac{1}{2^{n+1}}$. Hence, f_i converges uniformly to a continuous function called the cantor function defined as

$$f(x) := \lim_{i \to \infty} f_i(x)$$

and satisfy

$$|f(x) - f_n(x)| \leq \frac{1}{2^n}$$

.

Lemma 2.44. *The cantor function f defined on C satisfies the Hölder condition with the exponent $\gamma = \frac{log2}{log3}$.*

Proof. Note that f_n is piecewise continuous linear function and hence it is absolute continuous with the existence of $f_n'(x)$ a.e. Also, observe that $f_n'(x)$ is bounded by $\left(\frac{3}{2}\right)^n$. Therefore, for $x, y \in [0,1]$,

$$|f_n(x) - f_n(y)| = \left| \int_x^y f_n'(t)dt \right| \leq \left(\frac{3}{2}\right)^n |x - y|$$

Let $x \neq y$ and choose $n \geq 1$ such that $\frac{1}{3^n} \leq |x - y| \leq \frac{1}{3^{n-1}}$.

$$\implies |f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$$

$$\leq \frac{2}{2^n} + \left(\frac{3}{2}\right)^n |x - y|$$

$$\leq \frac{2}{2^n} + \frac{3}{2^n}$$

$$= \frac{5}{2^n}$$

$$\leq 5\left(\frac{1}{3^n}\right)^\gamma \leq 5|x - y|^\gamma$$

where $\gamma = s \equiv \frac{log2}{log3}$ $\square$

For the other inequality in the calculation of the dimension of the cantor set, let $E = C$, f be the cantor function defined as above and $s = \gamma = \frac{log2}{log3}$. The above 2 lemmas give $\mathcal{H}^1\big([0,1]\big) \leq M'\mathcal{H}^s(C)$. Thus $\mathcal{H}^s(C) > 0 \implies dim(C) \geq \frac{log2}{log3}$. Thus

$$dim(C) = \frac{log2}{log3}$$

$\square$

- **Hausdorff dimension of a general fractal :**

  Theory :

  Consider (X, d) to be a metric space. Let $A, B \subset X$, $x \in X$, $\delta > 0$.

  Define:

$$d(x, A) := \inf\{d(x, a) \ ; \ a \in A\}$$

$$A^\delta := \{x \in X \ ; \ d(x, A) < \delta\}$$

$$d(A, B) := \inf\{\delta \ ; \ A \subset B^\delta \ , \ B \subset A^\delta\}$$

Lemma (*1) : Let A, B, C be closed subsets of X. Then,

– $d(A, B) = d(B, A)$

– $d(A, B) = 0 \iff A = B$

– $d(A, B) \leq d(A, C) + d(C, B)$

– $A \subset B \implies A^\delta \subset b^\delta$

Proof. From the definition, $d(A, B) = d(B, A)$.

Clearly, $A = B \implies d(A, B) = 0$. Conversely, if $d(A, B) = 0$, then for all $\delta > 0$, $A \subset B^\delta$ and $B \subset A^\delta$. Hence, for every $a \in A$, there is a $b_\delta \in B$ such that $d(a, b_\delta) < \delta$. Let $\delta = \frac{1}{n}$, $b_n = b_\delta$. Then, $b_n \to a$ as $n \to \infty$. Hence, $a \in \overline{B} = B$ and thus $A \subset B$. Similarly, $B \subset A$ and this concludes that $A = B$.

Let $\delta_1 > 0$ and $\delta_2 > 0$ such that $d(A, C) < \delta_1$ and $d(B, C) < \delta_2$. Let $a \in A$, then there is a $c_1 \in C$ and $b_1 \in B$ such that $d(a, c_1) < \delta_1$ and $d(c_1, b_1) < \delta_2$.

$$\implies d(a, B) \leq d(a, b_1) \leq d(a, c_1) + d(c_1, b_1) < \delta_1 + \delta_2$$

$$\implies A \subset B^{\delta_1 + \delta_2}$$

Similarly, $B \subset A^{\delta_1 + \delta_2}$.

$$\implies d(A, B) \leq \delta_1 + \delta_2$$

Letting $\delta_1 \to d(A, C)$ and $\delta_2 \to d(B, c)$, we have the third point of the lemma proved.

Let $x \in A^\delta$. Then, $d(x, A) < \delta$ and hence, $d(x, B) < \delta$. thus $x \in B^\delta$ and hence,

$A^\delta \subset B^\delta$.

This proves the lemma.

$\square$

Definition:

Let $S : X \to X$ be a map. S is said to be a similarity map with ratio $0 < r < 1$, if $d(S(x), S(y)) = r\ d(x, y)$, for all $x, y \in X$.

Definition :

Let $S_1, S_2, \ldots, S_l$ be similarity maps with the same ratio r. Let $A \subset X$. Define,

$$L(A) := S_1(A) \cup \cdots \cup S_l(A)$$

Definition :

$A \subset X$ is said to be self similar with respect to $\{S_1, \ldots, S_l\}$ if $L(A) = A$.

Lemma (*2):

Let $S_1, S_2, \ldots, S_l$ be similarity maps on X with the same similarity constant. Let A, B be subsets of X and for $k \geq 0$, define

$$F_0(A) := A$$

$$F_k(A) := L\big(F_{k-1}(A)\big)$$

$$I(k,l) := \Big\{ J\ ;\ J : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, l\} \Big\}$$

Then,

- $diam(S(A)) \leq r\ diam(A)$
- For $J \in I(k,l)$,

$$diam\bigg(S_{J(1)} \circ \cdots \circ S_{J(k)}(A)\bigg) \leq r^k\ diam(A)$$

- For $J \in I(k,l)$, denote

$$S_J := S_{J(1)} \circ S_{J(2)} \circ \cdots \circ S_{J(k)}$$

  Then

$$F_k(A) = \bigcup_{J \in I(k,l)} S_J(A)$$

-

$$dist\bigg(L(A), L(B)\bigg) \leq r\ dist(A, B)$$

Proof. Since, $d\big(S(x), S(y)\big) = r\ diam(A, B)$, we have that $diam\big(S(A)\big) \leq r\ diam(A)$. Repeat this k many times to get,

$$diam\big(S_J(A)\big) = diam\big(S_{J(1)} \circ S_{J(2)} \circ \cdots \circ S_{J(k)}(A)\big) \leq r\ diam\big(S_{J(2)} \circ S_{J(3)} \circ \cdots \circ S_{J(k)}(A)\big)$$

$$\leq r^k\ diam(A)$$

Now, for the third point of the lemma. This is true by induction on k. For $k = 1$, it is true by the definition. Assume that the statement is true upto $k - 1$. Then,

$$F_k(A) = S_1(F_{k-1}(A)) \cup \cdots \cup S_l(F_{k-1}(A))$$

$$= \bigcup_{i=1}^{l} \bigcup_{J \in I(k-1,l)} S_J(A)$$

$$= \bigcup_{J \in I(k,l)} S_J(A)$$

Let $\delta > d(A, B)$, then $A \subset B^\delta$ and $B \subset A^\delta$.
Claim 1: $S(A) \subset S(B)^{r\delta}$, $S(B) \subset S(A)^{r\delta}$.
Proof of the claim 1: For $x \in S(A)$, $y \in S(B)$, then there is $a \in A$, $b \in B$ such that $S(a) = x$ and $S(b) = y$. Hence,

$$d(x, y) = d\big(S(a), S(b)\big) = r\ d(a, b)$$

And hence,

$$d\big(x, S(B)\big) \leq r\ d(a, b)\ \text{, for all } x \in S(A)$$

Now, taking infimum over A and B, we get

$$d(x, S(B)) \leq r \inf_{a \in A, b \in B} d(a, b) = r\ d(A, B) < r\delta$$

Hence, $S(A) \subset S(B)^{r\delta}$ and similarly, the same kind of proof gives $S(B) \subset S(A)^{r\delta}$. From the claim 1, we have for $1 \leq i \leq l$, $S_i(A) \subset S_i(B)^{r\delta}$ and $S_i(B) \subset S_i(A)^{r\delta}$. Thus, from the lemma(*1) as mentioned before,

$$L(A) = S_1(A) \cup \cdots \cup S_l(A)$$

$$\subset S_1(B)^{r\delta} \cup \cdots \cup S_l(B)^{r\delta}$$

$$\subset \Big(S_1(B) \cup \cdots \cup S_l(B)\Big)^{r\delta}$$

$$= L(B)^{r\delta}$$

Similarly, we also have $L(B) \subset L(A)^{r\delta}$. Hence,

$$dist\Big(L(A), L(B)\Big) \leq r\ \delta\ ,\ \forall \delta > dist(A, B)$$

By letting $\delta \downarrow d(A, B)$, we get

$$dist\Big(L(A), L(B)\Big) \leq r\ dist(A, B)$$

This proves the lemma. $\square$

Lemma(*3):

Let $S_1, S_2, \dots, S_l$ be similarity maps with the same ratio $0 < r < 1$. Then,

- Let $A \subset X$ such that $S_i(A) \subset A$, for all $1 \leq i \leq l$. Then

$$F_{k+1}(A) \subset F_k(A)$$

- For $i \leq l$, let $S_i(A) \subset A$ and $F(A) = \cap_{k=1}^{\infty} F_k(A)$, Then,

$$L\Big(F(A)\Big) = F(A)$$

- Let $x_0 \in X$, $\alpha := \max\Big\{dist\Big(S_i(x_0), x_0\Big)\ ;\ 1 \leq i \leq l\Big\}$ and $R > \frac{\alpha}{1-r}$, then,

$$L\Big(B(x_0, R)\Big) \subset B(x_0, R)$$

Proof. Given, for all $1 \leq i \leq l$, $S_i(A) \subset A$. Thus, for all $J \in I(k, l)$,

$$S_J(A) = S_{J(1)} \circ \cdots \circ S_{J(k)}(A) \subset S_{J(1)} \circ \cdots \circ S_{J(k-1)}(A) \subset F_{k-1}(A)$$

Hence, $F_k(A) = \bigcup_{J \in I(k,l)} S_J(A) \subset F_{k-1}(A)$. Now,

$$L\big(F(A)\big) = \bigcap_{k=0}^{\infty} L\big(F_k(A)\big) = \bigcap_{k=0}^{\infty} F_{k+1}(A) = \bigcap_{k=1}^{\infty} F_k(A) = F(A)$$

Since, $F_{k+1}(A) \subset F_k(A)$.

Now, let $R > \frac{\alpha}{1-r}$ and $x \in B(x_0, R)$, then for $1 \leq i \leq l$,

$$d\big(S_i(x), x_0\big) \leq d\big(S_i(x), S_i(x_0)\big) + d\big(S_i(x_0), x_0\big) \leq r\ d(x, x_0) + \alpha \leq rR + \alpha \leq R$$

This proves the lemma. $\square$

Lemma(*4) :
Assume that there is $x_0 \in X$ and for all $R > 0$, $\overline{B(x_0, R)}$ is compact. Then, there is a unique invariant compact set in X for $S_1, S_2, \dots, S_l$.

Proof. Uniqueness : Let E, F be two compact invariant sets for the family $S_1, \dots, S_l$. Then, $L(E) = E$ and $L(F) = F$, and from the lemma (*2),

$$d(E, F) = d\big(L(E), L(F)\big) \leq r\ d(E, F)$$

Since $r < 1$, we have $d(E, F) = 0$ and hence, $E = F$.
Existence : Let $\alpha = \max\Big\{d\big(S_i(x_0), x_0\big)\ ;\ 1 \leq i \leq l\Big\}$ and $R \geq \frac{\alpha}{1-r}$ and $A = \overline{B(x_0, R)}$. Then, from the lemma (*3), we have $S_i(A) \subset A$, $F_{k+1}(A) \subset F_k(A)$ and $F_k(A)$ is compact for all $k \geq 0$. Hence, $F(A) \neq \phi$ and is an invariant set to the family $\{S_1, \dots, S_l\}$.

□

Lemma (*5):
Let $X = \mathbb{R}^n$ and d be the euclidean distance. Let $S_1, S_2, \dots S_l$ be similarity maps with the same constant $0 < r < 1$. Let F be the invariant set for $\{S_1, S_2, \dots, S_l\}$. Then, for $s = \frac{log\ l}{log\ 1/r}$, we have $\mathcal{H}^s(F) < \infty$.

Proof. Hypothesis of the lemma (*4) is satisfied with $x_0 = 0$. Hence, for $A = \overline{B(x_0, R)}$, $F_{k+1}(A) = L\big(F_k(A)\big)$, we have

$$F = F(A) = \bigcap_{k=0}^{\infty} F_k(A) \text{ to be the invariant set.}$$

From the lemma(*2), for all $J \in I(k, l)$, we have

$$\begin{aligned} diam\big(S_J(A)\big) &= diam\Bigg(S_{J(1)} \circ \dots \circ S_{J(k)}(A)\Bigg) \\ &\leq r\ diam\Bigg(S_{J(2)} \circ \dots \circ S_{J(k)}(A)\Bigg) \\ &\leq r^k\ diam(A) \\ &= 2Rr^k \end{aligned}$$

Since, $F = F(A) \subset F_k(A) = \cup_{J \in I(k,l)} S_J(A)$, let $\delta \leq 2Rr^k$, then, as $lr^s = 1$, i.e, $s = \frac{log\ l}{log\ \frac{1}{r}}$,

$$\mathcal{H}^s_\delta\big(F(A)\big) \leq \alpha(s) \sum_{J \in I(k,l)} \left(\frac{diam(S_J(A))}{2}\right)^s$$

$$\le \alpha(s)\big(Rr^k\big)^s \; card\big(I(k,l)\big)$$

$$= \alpha(s)\big(Rr^k\big)^s l^k$$

$$= \alpha(s)R^s\big(lr^s\big)^k$$

$$\le \alpha(s)R^s$$

Since, $r < 1$, as $k \to \infty$, $\delta \to 0$ and therefore,

$$\mathcal{H}^s(F) = \mathcal{H}^s(F(A)) \le \alpha(s)R^s < \infty$$

This proves the lemma. □

Definition :

Let $S_1, S_2, \ldots, S_l$ be similarities with the same ratio r. Then, $S_1, S_2, \ldots, S_l$ are seperated if there is an open bounded set O such that

$$S_i(O) \cap S_j(O) = \phi \quad \text{if } i \neq j$$

$$\bigcup_{i=1}^{l} S_i(O) \subset O$$

Remark: Let $S_1, S_2, \ldots, S_l$ be seperated as above with the existence of O. Then, for any $J, K \in I(k,l)$, $J \neq K$,

- $S_J(O) \subset O$
- $S_J(O) \cap S_K(O) = \phi$

Proof. Since $J \neq K$, choose l_0 such that

$$J(i) = K(i) \text{ for all } i \le i_0 - 1$$

$$J(i_0) \neq K(i_0)$$

Since $S_i \circ S_j(O) \subset S_i(O)$,

$$S_J(O) = \Big(S_{J(1)} \circ \cdots \circ S_{J(i_0)} \circ \cdots \circ S_{J(k)}\Big)(O) \subset \Big(S_{J(1)} \circ \cdots \circ S_{J(i_0)}\Big)(O)$$

Similarly,

$$S_K(O) \subset S_{K(1)} \circ \cdots \circ S_{K(i_0)}(O)$$

Claim :

$$\left(S_{J(1)} \circ \cdots \circ S_{J(i_0)}(O)\right) \cap \left(S_{K(1)} \circ \cdots \circ S_{K(i_0)}(O)\right) = \phi$$

Proof of the claim : For, if x is in the intersection, then there is z_1 and z_2 such that

$$S_{J(1)} \circ \cdots \circ S_{J(i_0)}(z_1) = x = S_{K(1)} \circ \cdots \circ S_{K(i_0)}(O)(z_2)$$

Hence, by the choice of i_0, we have

$$S_{J(i_0)}(z_1) = S_{K(i_0)}(z_2)$$

This is clearly a contradiction and hence the claim is proved. Hence,

$$S_J(O) \cap S_K(O) = \phi \text{ for } J \neq K$$

This proves the remark. □

Lemma(*6) :
Let $A \subset \mathbb{R}^n$ be a compact set and $0 < s \leq n$ with $\mathcal{H}^s(A) < \infty$. Then,

$$\mathcal{H}^s_\delta(A) = \inf\left\{\sum_{j=1}^p \left(\frac{diam\ U_j}{2}\right)^s \ ;\ diam(U_j) < \delta\ ,\ A \subset \bigcup_{j=1}^p U_j\right\}$$

Proof. Since $\mathcal{H}^s(A) < \infty$, we have that $\mathcal{H}^s_\delta(A) < \infty$ for all $\delta > 0$.
Let $\epsilon > 0$, $0 < \delta < 1$ and $A \subset \bigcup_{j=1}^\infty C_j$, with C_j's closed and $diam(C_j) < \delta$ with the property that

$$\sum_{j=1}^\infty \alpha(s)\left(\frac{diam\ C_j}{2}\right)^s < \mathcal{H}^s_\delta(A) + \epsilon$$

Choose $\epsilon_j > 0$ such that $0 < \epsilon_j < diam(C_j)$, $diam(C_j) + 2\epsilon_j < 2\delta$ *and* $\sum_{j=1}^\infty \epsilon_j^{\min\{1,s\}} < \epsilon$. Then, for some $\lambda > 0$ which depends on s, we have

$$\sum_{j=1}^\infty \alpha(s)\left(\frac{diam(C_j)}{2} + \epsilon_j\right)^s \leq \sum_{j=1}^\infty \alpha(s)\left(\frac{diam\ C_j}{2}\right)^s + \lambda\sum_{j=1}^\infty \alpha(s)\epsilon_j^{\min\{1,s\}}$$

$$\leq \mathcal{H}^s_\delta(A) + \epsilon + \lambda\alpha(s)\epsilon$$

$$= \mathcal{H}^s_\delta(A) + (1 + \lambda\alpha(s))\epsilon$$

Let

$$U_j := \{x\ ;\ d(x, C_j) < \epsilon_j\}$$

$$\Longrightarrow \ diam(U_j) \leq diam(C_j) + 2\epsilon_j < \delta \ , \ A \subset \bigcup_{j=1}^{\infty} C_j \subset \bigcup_{j=1}^{\infty} U_j$$

Now, as A is compact, there is N such that $A \subset \bigcup_{j=1}^{N} U_j$. Then,

$$\sum_{j=1}^{N} \alpha(s) \left(\frac{diam\ U_j}{2} \right)^s \leq \sum_{j=1}^{\infty} \alpha(s) \left(\frac{diam\ C_j}{2} + \epsilon_j \right)^s = \mathcal{H}^s_\delta(A) + \big(1 + \lambda\alpha(s)\big)\epsilon$$

This proves the lemma as $\epsilon > 0$ was arbitrary. □

Let $\{S_1, S_2, \ldots, S_l\}$ be a seperating family of similarity transformations with the same ratio $0 < r < 1$, with the bounded open set to be O, as in the definition of the seperability. Let $R_0 > 0$ and $F \subset B(0, R_0)$ be the invariant set. Let $\delta < 1$ and $F \subset \bigcup_{i=1}^{p} U_i$ with $diam\ U_i < \delta$. Since $F \subset B(0, R_0)$, we can assume that $U_i \subset B(0, R_0)$ for all i. Let

$$\Lambda := \sup \Big\{ |v - z| \ ; \ v \in O \ , \ z \in B(0, R_0) \Big\}$$

Let $k \geq 1$ be such that

$$r^k \leq \min\{diam(U_i)\} < r^{k-1}$$

Let $\overline{x} \in F$ and $1 \leq p \leq k$.

Define :

$$\mathcal{G}_p := \{U_i \ ; \ r^p \leq diam(U_i) < r^{p-1}\}$$

$$\mathcal{G}_{p,i} := \left\{ \begin{array}{rr} J \equiv \big(J(1), \ldots, J(q)\big) \ ; & 1 \leq J(t) \leq l \ , \ \forall\ 1 \leq t \leq q \ , \ \forall\ q \geq l-1 \\ & S_J(\overline{x}) \in U_i \ , \ U_i \in \mathcal{G}_p \end{array} \right\}$$

Lemma (*7):

With the above definitions valid, there is a $c_0 > 0$ that depend on n, R_0, O such that

$$card(\mathcal{G}_{p,i}) \leq c_0 l^{k-p}$$

Proof. For $J \in \mathcal{G}_{p,i}$, declare $\overline{J} := \big(J(1), J(2), \ldots, J(p-1)\big)$.

Claim : There is a $c_0 > 0$ such that it depends on n, R_0, O and

$$card\{\overline{J} \ ; \ J \in \mathcal{G}_{p,i}\} \leq c_0$$

Proof of the claim : Let $J \in \mathcal{G}_{p,i}$. Then, for $v \in O$,

$$|J(\overline{x}) - \overline{J}(v)| \leq r^{p-1} \big| S_{J(p)} \circ \cdots \circ S_{J(q)}(\overline{x}) - v \big| \leq \Lambda r^{p-1}$$

Since, $J(\overline{x}) \in U_i$ and $diam(U_i) \leq r^{p-1}$, we have, for all $v \in O$ and for all $z \in U_i$,

$$|\overline{J}(v) - z| \leq |\overline{J}(v) - J(\overline{x})| + |J(\overline{x}) - z|$$

$$\leq \Lambda r^{p-1} + diam(U_i)$$

$$\leq (\Lambda + 1) r^{p-1}$$

$$\implies \overline{J}(O) \subset B\big(z, (\Lambda + 2) r^{p-1}\big)$$

Since $diam\big(\overline{J}(O)\big) = r^{p-1} diam(O)$ and $\{\overline{J}(O)\}$ are disjoint,
if $\mathcal{T} = card\{\overline{J}(O) \ ; \ J \in \mathcal{G}_{p,i}\}$, then,

$$\mathcal{T} \mathcal{L}^n(O) r^{n(l-1)} \leq (\Lambda + 2)^n \mathcal{L}^n\big(B(0,1)\big) r^{n(l-1)}$$

$$\implies \mathcal{T} \leq c_0 \ with \ c_o = \frac{(\Lambda + 2)^n \mathcal{L}^n\big(B(0,1)\big)}{\mathcal{L}^n(O)}$$

For each $\overline{J}$, there are atmost l^{k-p} elements in $\mathcal{G}_{p,i}$ and hence the cardinality of $\mathcal{G}_{p,i}$ is atmost $c_0 l^{k-p}$. This proves the lemma. □

Theorem :
Let $\{S_1, S_2, \ldots, S_l\}$ be a seperating family of similarity transformations with the common ratio to be $0 < r < 1$. Let $s = \frac{log \ l}{log \ 1/r}$ and F be the invariant set. Then, there exist $\alpha, \beta > 0$ such that $\alpha \leq \mathcal{H}^s(F) \leq \beta$. This concludes that the hausdorff dimension of F is $\frac{log \ l}{log \ 1/r}$.

Proof. From the lemma (*5), we have that $\mathcal{H}^s(F) \leq \beta$ for some $\beta > 0$. Also, from the lemmas (*6) and (*7), we prove the lower bound for $\mathcal{H}^s(F)$. Let $0 < \delta < 1$ and

$$F \subset \bigcup_{i=1}^{N} U_i \ ; \text{with } diam(U_i) < \delta.$$

Let $k \geq 1$ be such that

$$r^k \leq \min\{diam(U_i)\} < r^{k-1}$$

For $1 \leq p \leq k$, define

$$A(p) := \{i \ ; \ r^p \leq diam(U_i) < r^{p-1}\}$$

$$N(p) := card\Big(A(p)\Big)$$

Then, from the lemma (*7), there is a constant $c_0 > 0$ such that it is dependent on O and R_0 such that the following holds.

$$card\left(\bigcup_{p=1}^{k}\bigcup_{i\in A(p)}\mathcal{G}_{p,i}\right) \leq c_0\sum_{p=0}^{k-1}N(p)m^{k-p}$$

Since for $J \in I(k,l)$, $J(\overline{x}) \in F \subset \bigcup_{i=1}^{N}U_i$. And hence $J(\overline{x}) \in U_i$ for some i with $r^p \leq diam(U_i) < r^{p-1}$. Hence for $c_1 = \frac{1}{c_0}$,

$$\implies l^k \leq c_0\sum N(p)l^{k-p} \text{ or equivalently, } \sum_p \frac{N(p)}{l^p} \geq c_1$$

Now,

$$\begin{aligned}\sum_{i=1}^{N}\alpha(s)\left(\frac{diam(U_i)}{2}\right)^s &\geq \alpha(s)\sum_p N(p)r^{ps} \\ &= \alpha(s)\sum_p \frac{N(p)}{l^p}(r^s l)^p \\ &\geq \alpha(s)\sum_p \frac{N(p)}{l^p} \\ &\geq c_1\end{aligned}$$

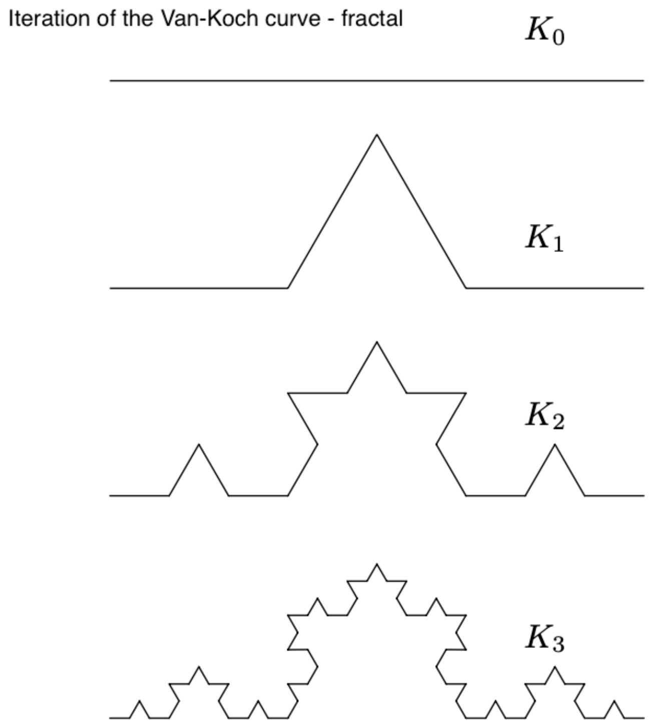
Hence,

$$\mathcal{H}^s(F) = \lim_{\delta\to 0}\mathcal{H}^s_\delta(F) \geq c_1$$

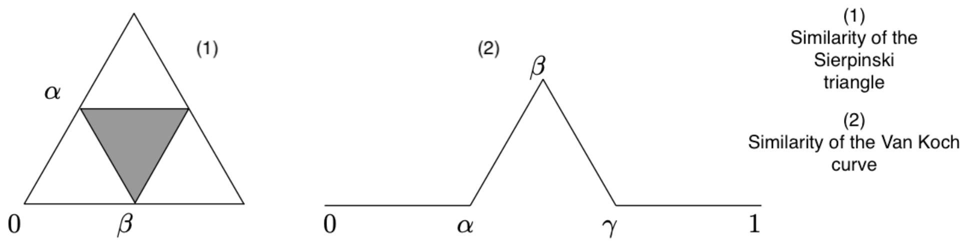
This proves the theorem.

□

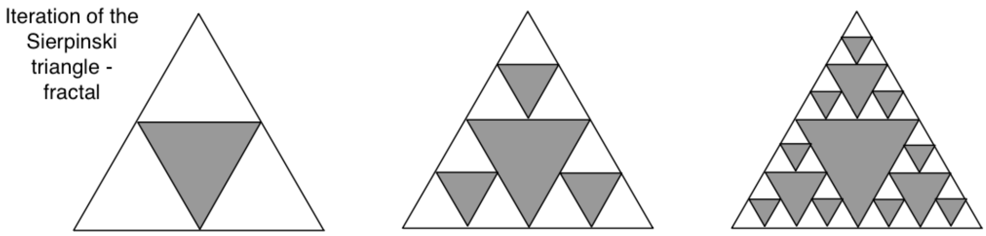
- **Examples as facts with the application of the previous theorem :**
  As seen earlier, the cantor set has dimension $\frac{log2}{log3}$.
  The van Koch curve has dimension $\frac{log4}{log3}$, where the similarities and the details are given by
  - $S_1(x) = \frac{x}{3}$
  - $S_2(x) = \rho\frac{x}{3} + \alpha$
  - $S_3(x) = \frac{1}{\rho}\frac{x}{3} + \beta$
  - $S_4(x) = \frac{x}{3} + \gamma$
  - Here ρ is the rotation centered at the origin and of angle $\frac{\pi}{3}$, that is (say) $e^{i\cdot\frac{\pi}{3}}$
  - Here, $m = 4, r = 1/3$
  - Refer the diagram for the fractal as well as the definition of α, β, γ.



Fig(a) : Figures of the iteration of the Van Koch curve - fractal.



Fig(b) : Figures of the similarities of both the van Koch curve and the Sierpinski triangle.



Fig(c) : Figures of the iteration of the Sierpinski triangle - fractal.

The Sierpinski triangle has dimension $\frac{log3}{log2}$. The similarities and the details are given by

- $S_1(x) = \frac{x}{2}$
- $S_2(x) = \frac{x}{2} + \alpha$
- $S_3(x) = \frac{x}{2} + \beta$
- Here, $m = 3$ and $r = \frac{1}{2}$. Refer the diagram for the fractal and the definition of α and β which are basically midpoints of the sides on the triangle as shown above.

□

2.3.3 Isodiametric Inequality

Goal : To show that $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$.

Lemma 2.45. *Let $f : \mathbb{R}^n \to [0, \infty)$ be $\mathcal{L}^n$ measurable. Then, "the region under the graph of f" which is*

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

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

Proof. Let $g(x, y) := f(x) - y$. Note that, by setting $h(x, y) = f(x)$ and $\alpha(x, y) = y$, for any $\beta \in \mathbb{R}$,

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\ ;\ h(x, y) < \beta\} = \{x \in \mathbb{R}^n\ ;\ f(x) < \beta\} \times \mathbb{R}$$

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}\ ;\ \alpha(x, y) = y < \beta\} = \mathbb{R}^n \times \{y \in \mathbb{R}\ ;\ y < \beta\} = \mathbb{R}^n \times (-\infty, \beta)$$

Both the above sets belong to the $\mathcal{L}^{n+1} - \sigma-$ algebra and thus g is $\mathcal{L}^{n+1}$ measurable and

$$A = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}\ ;\ y \geq 0\} \cap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}\ ;\ g(x, y) \geq 0\}$$

Thus A is $\mathcal{L}^{n+1}$ measurable. □

Definition 2.46. Define for $a, b \in \mathbb{R}^n$,

$$L_b^a := \{b + t.a\ ;\ t \in \mathbb{R}\} \equiv \text{The line through } b \text{ in the direction of } a.$$

$$P_a := \{x \in \mathbb{R}^n\ ;\ x.a = 0\} \equiv \text{Plane through the origin perpendicular to } a.$$

Definition 2.47. Steiner Symmetrization of A.
Choose $a \in \mathbb{R}^n$ with $||a|| = 1$. Let $A \subset \mathbb{R}^n$. We define the steiner symmetrization of A with respect to P_a to be the set

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

Lemma 2.48. *Properties of Steiner symmetrization :*

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

Proof. Part (a) :
Case 1 : If $diam(A) = \infty$, then the first part is trivial in conclusion.
Case 2 : $diam(A) < \infty$.
Since, the diameter is independent of the closures, we may assume that A is closed. Fix $\epsilon > 0$ and let $x, y \in S_a(A)$ such that by the supremum property,

$$diam(S_a(A)) \leq |x - y| + \epsilon$$

Set

$$b = x - (x.a)a$$

$$c = y - (y.a)a$$

Then, clearly,

$$b.a = 0 = c.a \implies b, c \in P_a$$

Set

$$r = \inf\{t \ ; \ b + ta \in A\}$$

$$s = \sup\{t \ ; \ b + ta \in A\}$$

$$u = \inf\{t \ ; \ c + ta \in A\}$$

$$v = \sup\{t \ ; \ c + ta \in A\}$$

Note that all the 4 above defined quantities are finite as $diam(A) < \infty$. W.L.O.G assume that $v - r \geq s - u$. Then,

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

The above estimate is true because

$$A \cap L_b^a \subset I = \{b + ta \ ; \ |t| \le \frac{s-r}{2}\}$$

$$\implies \mathcal{H}^1\big(A \cap L_b^a\big) \le \mathcal{H}^1(I) = \mathcal{H}^1\big(\{t.e_1 \ ; \ |t| \le \frac{s-r}{2}\}\big) = \mathcal{L}^1\left[\frac{r-s}{2}, \frac{s-r}{2}\right] \equiv s-r$$

Now, $b + (x.a)a \in S_a(A)$ and $c + (y.a)a \in S_a(A)$ and hence,

$$|x.a| \le \frac{1}{2}\mathcal{H}^1(A \cap L_b^a)$$

$$|y.a| \le \frac{1}{2}\mathcal{H}^1(A \cap L_c^a)$$

Thus

$$v - r \ge \frac{1}{2}\mathcal{H}^1\big(A \cap L_b^a\big) + \frac{1}{2}\mathcal{H}^1\big(A \cap L_c^a\big) \ge |x.a| + |y.a| \ge |x.a - y.a|$$

Therefore,

$$(diam(S_a(A) - \epsilon))^2 \le |x-y|^2 = |b-c|^2 + |x.a - y.a|^2$$

$$\le |b-c|^2 + |v-r|^2 = |(b+ra) - (c+va)|^2 \le diam(A)^2$$

This concludes that

$$diam(S_a(A)) \le diam(A)$$

Part (b) :

Observe that $\mathcal{L}^n$ is rotation invariant and thus assume $a = e_n \equiv (0, 0, \ldots, 1)$ and $P_a = P_{e_n} \equiv \mathbb{R}^{n-1}$. Now $\mathcal{L}^1 = \mathcal{H}^1$ on $\mathbb{R}$.

A is $\mathcal{L}^n$ measurable implies

$$\mathcal{L}^n(A) = \int_{\mathbb{R}^n} \mathcal{X}_A$$

$$= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \mathcal{X}_A(x_1, x_2, \ldots, x_n) dx_n \right) dx_2 \ldots . dx_{n-1}$$

$$= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \mathcal{X}_A(b, x_n) dx_n \right) db$$

Notice that fixing b, x_n can be on the line joining b, e_n in A, and thus

$$= \int_{\mathbb{R}^{n-1}} \mathcal{L}^1(\mathcal{X}_{A \cap L_b^a}) db$$

$$\implies f(b) = \mathcal{L}^1(\mathcal{X}_{A \cap L_b^a}) = \mathcal{H}^1(\mathcal{X}_{A \cap L_b^a}) \text{ is } \mathcal{L}^{n-1} \text{ measurable a.e.}$$

Now,

$$S_a(A) = \{(b, y) \ ; \ \frac{-f(b)}{2} \le y \le \frac{f(b)}{2}\} - \{(b, 0) \ ; \ L_b^a \cap A = \phi\}$$

By the previous lemma, $\{(b,y)\ ;\ \frac{-f(b)}{2} \leq y \leq \frac{f(b)}{2}\}$ is $\mathcal{L}^n$ measurable and $\{(b,0)\ ;\ L_b^a \cap A = \phi\} \subset \mathbb{R}^{n-1} \times \{0\}$, which has measure 0, in $\mathcal{L}^n$ measure. Thus $S_a(A)$ is $\mathcal{L}^n$ measurable and so

$$\mathcal{L}^n(S_a(A)) = \mathcal{L}^n\{(b,y)\ ;\ |y| \leq \frac{f(b)}{2}\}+0 = \int_{\mathbb{R}^{n-1}} \left(\int_{\frac{-f(b)}{2}}^{\frac{f(b)}{2}} dy \right) db = \int_{\mathbb{R}^{n-1}} f(b)db = \mathcal{L}^n(A)$$

$\square$

Theorem 2.49. *Isodiametric inequality :*
For all sets $A \subset \mathbb{R}^n$,

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

Remark 2.50. Note that A need not be contained in a ball of diameter $\equiv diam(A)$.

Proof. Case 1 : $diam(A) = \infty$, then there is nothing more to be shown.
Case 2 : $diam(A) < \infty$.
Let $e_1, e_2, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. Define

$$A_1 = S_{e_1}(A)$$

$$A_2 = S_{e_2}(A_1)$$

$$\vdots$$

$$A^* \equiv A_n = S_{e_n}(A_{n-1})$$

Claim 1 : A^* is symmetric around the origin.
Proof of the claim 1 : A_1 is symmetric with respect to P_{e_1} . As

$$S_{e_1}(A) = \bigcup_{\{b \in P_{e_1}\ ;\ A \cap L_b^{e_1} \neq \phi\}} \{b + te_1\ ;\ |t| \leq \frac{1}{2}\mathcal{H}^1(A \cap L_b^{e_1})\}$$

Let $x \equiv (x_1, x_2, \ldots, x_n) \in S_{e_1}(A) \implies x = b + te_1$, for some $b \in P_{e_1}$ and $A \cap L_b^{e_1} \neq \phi$ and $|t| \leq \mathcal{H}^1(A \cap L_b^{e_1})$. Noting that $b_1 = 0$, $x \equiv (t, b_2, \ldots, b_n)$. Now, the reflection of $x \equiv (x_1, x_2, \ldots x_n)$ around e_1 would be $(-x_1, x_2, \ldots x_n)$, i.e $(-t, b_2, \ldots, b_n)$. Since $|t| \leq \frac{1}{2}\mathcal{H}^1\big(A \cap L_b^{e_1}\big)$ and hence $(-t, b_2, \ldots, b_b) \in S_{e_1}(A)$.
Now, let $1 \leq k \leq n$ and suppose that A_k is symmetric with respect to $P_{e_1}, \ldots, P_{e_k}$.
It is required to show that A_{k+1} is symmetric with respect to $P_{e_1}, \ldots P_{e_{k+1}}$.

$$A_{k+1} \equiv S_{e_{k+1}}(A_k) = \bigcup_{\{b \in P_{e_{k+1}}\ ;\ A_k \cap L_b^{e_{k+1}} \neq \phi\}} \{b + te_{k+1}\ ;\ |t| \leq \frac{1}{2}\mathcal{H}^1(A_k \cap L_b^{e_{k+1}})\}$$

By the similar logic that A_1 is symmetric with P_{e_1}, A_{k+1} is symmetric with respect to $P_{e_{k+1}}$. Now, fix $1 \leq j \leq k$. Let $S_j : \mathbb{R}^n \to \mathbb{R}^n$ be the reflection around P_j. Let $b \in P_{k+1}$. By the induction hypothesis,

$$S_j(A_k) = A_k$$

Here, the fact that reflection is an isometry is getting used and that hausdorff measures are invariant under the isometries. Thus

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

Thus

$$S_j(A_{k+1}) = A_{k+1}$$

Thus A_{k+1} is symmetric with respect to P_{e_j}. And, hence, $A^* = A_n$ is symmetric with respect to $P_{e_1}, P_{e_2}, \ldots P_{e_n}$.
This proves the claim 1.

Claim 2: $\mathcal{L}^n(A^*) \leq \alpha(n) \left(\frac{diam(A^*)}{2}\right)^2$.
Proof of the claim 2 : By the claim 1, $x \in A^* \iff -x \in A^*$. So $diam(A^*) \geq 2|x|$ and thus

$$A^* \subset B\left(0, \frac{diam(A^*)}{2}\right)$$

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

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

This proves the claim 2.

Claim 3 : $\mathcal{L}^n(A) \leq \alpha(n) \left(\frac{diam(A)}{2}\right)^n$.
Proof of the claim 3 : Observe that if A is $\mathcal{L}^n$ measurable, then $\overline{A}$ is also $\mathcal{L}^n$ measurable. The previous lemma, thus tells that

- $\mathcal{L}^n((\overline{A})^*) = \mathcal{L}^n(\overline{A})$
- $diam((\overline{A})^*) \leq diam(\overline{A})$

Hence,

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

$$\leq \alpha(n)\left(\frac{diam(\overline{A})^*}{2}\right)^n$$

$$\leq \alpha(n)\left(\frac{diam(\overline{A})}{2}\right)^n$$
$$= \alpha(n)\left(\frac{diam(A)}{2}\right)^n$$

This proves the claim 3 and hence, the isodiamteric inequality. □

Theorem 2.51. *$\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$.*

Proof. Claim 1 : For all $A \subset \mathbb{R}^n$, $\mathcal{L}^n(A) \leq \mathcal{H}^n(A)$.
Proof of the claim 1 : Fix $\delta > 0$. Choose sets $\{C_j\}_{j=1}^{\infty}$ such that

- $A \subset \bigcup C_j$.
- $diam(C_j) \leq \delta$.

By the isodiametric inequality,

$$\mathcal{L}^n(A) \leq \sum_j \mathcal{L}^n(C_j) \leq \sum_j \alpha(n)\left(\frac{diam(C_j)}{2}\right)^n$$

Taking infima over the values, we get

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

Note that δ was arbitrary and hence,

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

This proves the claim 1.

By the definition of $\mathcal{L}^n \equiv \mathcal{L}^1 \times \ldots_{\text{n times}} \times \mathcal{L}^1$, for $A \subset \mathbb{R}^n$ and $\delta > 0$,

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

Claim 2 : $\mathcal{H}^n$ is absolutely continuous with repsect to $\mathcal{L}^n$.
Proof of the claim 2 : Set

$$C_n := \alpha(n)\left(\frac{\sqrt{n}}{2}\right)^n$$

For all cubes $Q \subset \mathbb{R}^n$,

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

$$\implies C_n\mathcal{L}^n(Q) = \alpha(n)\left(\frac{diam(Q)}{2}\right)^n$$

Thus,

$$\mathcal{H}^n_\delta(A) \leq \inf\{\sum_{i=1}^{\infty}\alpha(n)\left(\frac{diam(Q_i)}{2}\right)^n \ ; \ Q_i \text{ are cubes }, A \subset \bigcup_{i=1}^{\infty} Q_i \ , \ diam(Q_i) < \delta\}$$

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

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

Let $\delta \to 0$, to get

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

Thus, $\mathcal{L}^n(A) = 0 \implies \mathcal{H}^n_\delta(A) \leq 0 \implies \mathcal{H}^n(A) = 0$.
This proves the claim 2.

Claim 3 : $\mathcal{H}^n(A) \leq \mathcal{L}^n(A)$, for all $A \subset \mathbb{R}^n$.
Proof of the claim 3 : Fix $\delta, \epsilon > 0$ and select cubes $\{Q_i\}_{i=1}^{\infty}$ such that $A \subset \bigcup Q_i$, $diam(Q_i) < \delta$ and

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

By the second corollary, following the vitalli covering lemma, there is $\{B^i_k\}_{k=1}^{\infty}$, disjoint closed balls in Q_i° such that

$$diam(B^i_k) \leq \delta$$

$$\mathcal{L}^n\left(Q_i - \bigcup_{k=1}^{\infty} B^i_k\right) = \mathcal{L}^n\left(Q_i^\circ - \bigcup_{k=1}^{\infty} B^i_k\right) = 0$$

By the claim 2,

$$\mathcal{H}^n_\delta\left(Q_i^\circ - \bigcup_{k=1}^{\infty} B^i_k\right) = 0$$

Thus

$$\mathcal{H}^n_\delta(A) \leq \sum_{i=1}^{\infty}\mathcal{H}^n_\delta(Q_i)$$

$$\leq \sum_{i=1}^{\infty}\mathcal{H}^n_\delta\left(\bigcup_{i=1}^{\infty} B^i_k\right)$$

$$\leq \sum_{i=1}^{\infty}\sum_{k=1}^{\infty}\mathcal{H}^n_\delta(B^i_k)$$

$$\leq \sum_{i,k=1}^{\infty} \alpha(n). \left(\frac{diam(B_k^i)}{2}\right)^n$$

$$= \sum_{i,k=1}^{\infty} \mathcal{L}^n(B_k^i)$$

$$= \sum_{i=1}^{\infty} \mathcal{L}^n \left(\bigcup_{k=1}^{\infty} B_k^i\right)$$

$$= \sum_{i=1}^{\infty} \mathcal{L}^n(Q_i)$$

$$\leq \mathcal{L}^n(A) + \epsilon$$

Thus,

$$\mathcal{H}_\delta^n(A) \leq \mathcal{L}^n(A) + \epsilon$$

Let $\delta \to 0$ and then $\epsilon \to 0$, we get

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

Along with the claim 1, we can conclude that

$$\mathcal{H}^n = \mathcal{L}^n \text{ on } \mathbb{R}^n$$

□

Remark 2.52. Assume for the rest of the type-up, unless mentioned that $0 < s < n$.

Theorem 2.53. *Assume $E \subset \mathbb{R}^n$, E is $\mathcal{H}^s$ measurable and $\mathcal{H}^s(E) < \infty$. Then,*

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

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

Proof. Fix $t > 0$. Define

$$A_t := \{x \notin E\ ;\ \limsup_{r \to 0} \frac{\mathcal{H}^s\Big(B(x,r) \cap E\Big)}{\alpha(s)\ r^s} > t\}$$

Want to show that $\mathcal{H}^s(A_t) = 0$, for all $t > 0$.
Now, E is $\mathcal{H}^s$ measurable and $\mathcal{H}^s(E) < \infty$ implies that $(\mathcal{H}^s|E)$ is a radon measure and finite.
So, given $\epsilon > 0$, there is a compact set $K \subset E$ such that $\mathcal{H}^s(E - K) < \epsilon$. Let $U = \mathbb{R}^n - K$

be an open set. This implies that as $\mathbb{R}^n - E \subset \mathbb{R}^n - K$, $A_t \subset \mathbb{R}^n - E \subset \mathbb{R}^n - K = U$. Fix $\delta > 0$ and consider

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

Clearly, $\mathcal{F}$ covers A_t. Hence, by the vitali covering theorem (2.2), there is countable disjoint family of balls $\{B_i\}_{i=1}^{\infty}$ such that

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

Let $B_i \equiv B(x_i, r_i)$, then,

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(A_t) &\leq \sum_{i=1}^{\infty} \alpha(s)(5r_i)^s \\
&\leq \frac{5^s}{t} \sum_{i=1}^{\infty} \mathcal{H}^s(B_i \cap E) \\
&\leq \frac{5^s}{t} \mathcal{H}^s(U \cap E) \\
&= \frac{5^s}{t} \mathcal{H}^s(E - K) \\
&\leq \frac{5^s}{t} \epsilon
\end{aligned}$$

Noting that $\delta > 0$ was arbitrary,

$$\implies \mathcal{H}^s(A_t) \leq \frac{5^s}{t}\epsilon$$

Note, that if the proof was started with $t = 0$, then the previous step would not have been justified in this proof. Let $\epsilon \to 0$,

$$\implies \mathcal{H}^s(A_t) = 0$$

$\square$

Theorem 2.54. *Assume $E \subset \mathbb{R}^n$ and is $\mathcal{H}^s$ measurable. Also it is given that $\mathcal{H}^s(E) < \infty$. Then,*

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

for all $\mathcal{H}^s$ almost every $x \in E$.

Proof. Claim 1 : $\limsup_{r\to 0} \frac{\mathcal{H}^s\left(B(x,r)\cap E\right)}{\alpha(s)\ r^s} \leq 1$ for $\mathcal{H}^s$ almost every $x \in E$.
Proof of the claim 1 : Fix $\epsilon > 0$ and $t > 1$. Define

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

Now, $(\mathcal{H}^s|E)$ is radon measure and finite. This implies there is an open set U such that

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

This implies that $\mathcal{H}^s(U\cap E) \leq \mathcal{H}^s(U) \leq \mathcal{H}^s(B_t) + \epsilon$. Define,

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

Clearly, $\mathcal{F}$ covers B_t. By the corollary (2.3) to the vitali covering theorem, there are disjoint balls $\{B_i\}$ in $\mathcal{F}$ such that for a fixed $m \in \mathbb{N}$,

$$B_t - \bigcup_{i=1}^{m} B_i \subset \bigcup_{i=m+1}^{\infty} \hat{B}_i$$

$$\implies B_t \subset \left(\bigcup_{i=1}^{m} B_i\right) \bigcup \left(\bigcup_{i=m+1}^{\infty} \hat{B}_i\right)$$

For $B_i \equiv B(x_i, r_i)$,

$$\mathcal{H}^s_{10\delta}(B_t) \leq \sum_{i=1}^{m} \alpha(s) r_i^s + \sum_{i=m+1}^{\infty} \alpha(s)(5r_i)^s$$

$$\leq \frac{1}{t}\sum_{i=1}^{m} \mathcal{H}^s(B_i\cap E) + \frac{5^s}{t}\sum_{i=m+1}^{\infty} \mathcal{H}^s(B_i \cap E)$$

$$\leq \frac{1}{t}\mathcal{H}^s(U\cap E) + \frac{5^s}{t}\mathcal{H}^s\left(\bigcup_{i=m+1}^{\infty} B_i \cap E\right)$$

The above is true for all m and hence as $\mathcal{H}^s(\cup B_i \cap E) \leq \mathcal{H}^s(E) < \infty$, taking $m \to \infty$ and by the convergence of the tail of convergent sequence,to 0, we have

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

Let $\delta \to 0$ and then $\epsilon \to 0$, then

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

Now, noting that $\mathcal{H}^s(B_t) \leq \mathcal{H}^s(E) < \infty$, and if $\mathcal{H}^s(B_t) \neq 0$, we have $1 \leq \frac{1}{t}$, which is a contradiction.
This proves the claim 1.

Claim 2 : $\limsup_{r\to 0} \frac{\mathcal{H}^s_\infty\left(B(x,r)\cap E\right)}{\alpha(s)\ r^s} \geq \frac{1}{2^s}$, for $\mathcal{H}^s$ almost every $x \in E$.
Proof of the claim 2 : Fix $\delta > 0$ and $0 < t < 1$. Define

$$E(\delta,t) := \left\{x \in E\ ;\ x \in C\ ,\ C \subset \mathbb{R}^n\ ,\ diam(C) < \delta \implies \mathcal{H}^s_\delta(C \cap E) \leq t\ \alpha(s) \left(\tfrac{diam(C)}{2}\right)^s\right\}$$

Therefore, if $\{C_i\}_{i=1}^\infty$ are subsets or $\mathbb{R}^n$ such that $diam(C_i) \leq \delta$, $E(\delta,t) \subset \bigcup_{i=1}^\infty C_i$, $C_i \cap E(\delta,t) \neq \phi$,

$$\mathcal{H}^s_\delta\big(E(\delta,t)\big) \leq \sum_{i=1}^\infty \mathcal{H}^s_\delta\left(C_i \cap E(\delta,t)\right) \leq \sum_{i=1}^\infty \mathcal{H}^s_\delta(C_i \cap E) \leq t \sum_{i=1}^\infty \alpha(s) \left(\frac{diam(C_i)}{2}\right)^s$$

By the infimum property,

$$\mathcal{H}^s_\delta(E(\delta,t)) \leq t\ \mathcal{H}^s_\delta(E(\delta,t))$$

By the same logic as before, since $\mathcal{H}^s_\delta(E(\delta,t)) \leq \mathcal{H}^s_\delta(E) \leq \mathcal{H}^s(E) < \infty$, if $\mathcal{H}^s_\delta(E(\delta,t)) \neq 0$, then $1 \leq t$, which is a contradiction.
In particular, $\mathcal{H}^s(E(\delta, 1-\delta)) = 0$.
Now, if $x \in E$ and $\limsup_{r\to 0} \frac{\mathcal{H}^s_\infty\left(B(x,r)\cap E\right)}{\alpha(s)\ r^s} < \frac{1}{2^s}$, then there is a $\delta > 0$ such that $\frac{\mathcal{H}^s_\infty\left(B(x,r)\cap E\right)}{\alpha(s)\ r^s} \leq \frac{1-\delta}{2^s}$ for all $0 < r < \delta$. Thus , if $x \in C$ and $diam(C) \leq \delta$,

$$\mathcal{H}^s_\delta(C \cap E) = \mathcal{H}^s_\infty(C \cap E) \leq \mathcal{H}^s_\infty\Big(B\big(x, diam(C)\big) \cap E\Big) \leq (1-\delta)\ \alpha(s) \left(\frac{diam(C)}{2}\right)^s$$

Thus, $x \in E(\delta, 1-\delta)$. But then

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

This proves the claim 2.
Now, $\mathcal{H}^s\Bigg(B(x,r) \cap E\Bigg) \geq \mathcal{H}^s_\infty\Bigg(B(x,r) \cap E\Bigg) \geq \frac{1}{2^s}$, $\mathcal{H}^s$ a.e.

$\square$

Remark 2.55. It is possible to have $\limsup < 1$ and $\liminf = 0$.
Consider $E = \{x\}$ and consider $\mathcal{H}^1$ on it.

2.3.4 Relations to lipshitz mappings

Definition 2.56.
- $f : \mathbb{R}^n \to \mathbb{R}^m$ is lipshitz if there is a constant $C > 0$ such that

$$\forall x, y \in \mathbb{R}^n \; ; \; |f(x) - f(y)| \leq C.|x - y|$$

-

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

- $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ is locally lipshitz if for all compact sets $K \subset A$, there is a constant depending on K , $C(K)$ such that

$$|f(x) - f(y)| \leq C(K)|x - y| \text{ , for all } x, y \in K$$

Theorem 2.57. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be lipshitz. $A \subset \mathbb{R}^n$. $0 \leq s < \infty$. Then,*

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

Proof. Fix $\delta > 0$. Choose $\{C_i\}_{i=1}^{\infty}$ such that $diam(C_i) \leq \delta$ and $A \subset \bigcup_{i=1}^{\infty} C_i$. Then,

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

$$f(A) \subset \bigcup_{i=1}^{\infty} f(C_i)$$

Thus,

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

And hence,

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

Letting $\delta \to 0$ gives that

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

□

Corollary 2.58. *Suppose that $n > k$. Let $P : \mathbb{R}^n \to \mathbb{R}^k$ be the usual projection, i.e $P(x_1, x_2, \ldots x_n) = (x_1, x_2, \ldots, x_k)$. Let $A \subset \mathbb{R}^n$. Let $0 \leq s < \infty$. Then,*

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

Proof. Projection maps are lipshitz with $Lip(f) \leq 1$, that is

$$\frac{\sum_{i=1}^{k} |x_i - y_i|}{\sum_{i=1}^{n} |x_i - y_i|} \leq 1$$

The previous theorem 2.57, now concludes the corollary. $\square$

Definition 2.59. For $f : \mathbb{R}^n \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, define the graph of f to be the set

$$G(f; A) := \{(x, f(x))\ x \in A\} \subset \mathbb{R}^n \times \mathbb{R}^m \equiv \mathbb{R}^{n+m}$$

Theorem 2.60. *Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ and $\mathcal{L}^n(A) > 0$. Then,*

- $\mathcal{H}_{dim}(G(f; A)) \geq n$
- *If f is lipshitz, then $\mathcal{H}_{dim}(G(f; a)) = n$.*

Proof.
- Let $P : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be a projection. Thus, by the corollary,

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

$$\implies \mathcal{H}_{dim}(G(f; A)) \geq n$$

- Let Q be a cube of length 1 in $\mathbb{R}^n$. Sub-divide Q into k^n many sub-cubes each having the length $\frac{1}{k}$. Call these cubes by the index $Q_1, Q_2, \dots Q_{k^n}$.
  Observe that the $diam(Q_i) = \frac{\sqrt{n}}{k}$. Let, for $f \equiv (f_1, f_2, \dots, f_m)$,

$$a_{ij} := \min_{x \in Q_j} f^i(x)$$

$$b_{ij} := \max_{x \in Q_j} f^i(x)$$

  As f is lipshitz, we have, for $a_j = (a_{1,j}, a_{2,j}, \dots, a_{n,j})$ and $b_j = (b_{1,j}, b_{2,j} \dots, b_{n,j})$,

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

  Now, let

$$C_j \equiv Q_j \times \prod_{i=1}^{m} (a_{ij}, b_{ij})$$

  Then,

$$\{\big(x, f(x)\big)\ ;\ x \in Q_j \cap A\} \subset C_j$$

And

$$diam(C_j) \le \frac{Lip(f)\sqrt{n}}{k}$$

Since, $G(f; A \cap Q) \subset \bigcup_{j=1}^{k^n} C_j$,

$$\mathcal{H}^n_{\sqrt{n}.Lip(f)/k}\Big(G(f; A \cap Q)\Big) \le \sum_{j=1}^{k^n}\left(\frac{diam(C_j)}{2}\right)^n \alpha(n) = \left(\frac{Lip(f)\sqrt{n}}{2k}\right)^n k^n\alpha(n)$$

$$= \left(\frac{Lip(f)\sqrt{n}}{2}\right)^n \alpha(n)$$

Now, let $k \to \infty$,

$$\implies \mathcal{H}^n(G(f; A \cap Q)) < \infty$$

$$\implies \mathcal{H}_{dim}(G(f; A \cap Q)) \le n$$

Since the above is true for all cubes in $\mathbb{R}^n$ of side length 1 and $\mathbb{R}^n$ can be covered by countable unit cubes, we have the conclusion for the theorem.

□

Theorem 2.61. *Let $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$ and suppose that $0 \le s < n$, define*

$$\Lambda_s := \{x \in \mathbb{R}^n \; ; \; \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f| \; dy > 0\}$$

Then

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

Proof. Assume that $f \in \mathcal{L}^1_{loc}(\mathbb{R}^n)$. Thus, for all compact sets $K \subset \mathbb{R}^n$, $\int_K |f| < \infty$. Declaring the value to be 0 outside that compact set, we can assume that $f \in \mathcal{L}^1(\mathbb{R}^n)$. By the Lebesgue Besicovitch differentiation theorem (2.18),

$$\lim_{r\to 0} \fint_{B(x,r)} |f| \; dy = |f(x)|$$

Thus,

$$\lim_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f| \; dy = 0$$

The above is true for $\mathcal{L}^n$ almost every x.
Hence, $\mathcal{L}^n(\Lambda_s) = 0$. Now fix $\epsilon, \delta, \sigma > 0, f \in \mathcal{L}^1 \implies \exists \eta > 0$ such that, for all $A \subset \mathbb{R}^n$, $\mathcal{L}^n(A) < \eta$ and $\int_A |f| \; dx < \sigma$. Define

$$\Lambda_s^\epsilon := \{x \in \mathbb{R}^n \; ; \; \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f| \; dy > \epsilon\}$$

Thus $\Lambda_s^\epsilon \subset \Lambda_s \implies \mathcal{L}^n(\Lambda_s^\epsilon) = 0$. Thus there is an open set U such that

- $\Lambda_s^\epsilon \subset U$
- $\mathcal{L}^n(U) < \eta$

Set

$$\mathcal{F} := \{B(x,r) \ ; \ x \in \Lambda_s^\epsilon \ , \ 0 < r < \delta \ , \ B(x,r) \subset U \ , \ \int_{B(x,r)} |f| \ dy > \epsilon \ r^s\}$$

Observe that $\mathcal{F}$ covers Λ_s^ϵ. Thus, by the vitali covering theorem, there are disjoint balls $\{B_i\}_{i=1}^\infty \subset \mathcal{F}$ such that

$$\Lambda_s^\epsilon \subset \bigcup_{i=1}^\infty \hat{B}_i$$

Let $B_i \equiv B(x_i, r_i)$,

$$\mathcal{H}^s_{10\delta}(\Lambda_s^\epsilon) \leq \sum_{i=1}^\infty \alpha(s)(5r_i)^s \leq \frac{\alpha(s)5^s}{\epsilon} \sum_{i=1}^\infty \int_{B_i} |f| \ dy \leq \frac{\alpha(s)5^s}{\epsilon} \int_U |f| \ dy \leq \frac{\alpha(s)5^s}{\epsilon} \ \sigma$$

Sending $\delta \to 0$ and the $\sigma \to 0$, we have

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

$\square$

2.3.5 Rademacher's Theorem.

Theorem 2.62. ***Extension theorem:***
Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be lipshitz. Then, there is $\overline{f} : \mathbb{R}^n \to \mathbb{R}^m$ such that

- $\overline{f} = f$ *on* A.
- $Lip(\overline{f}) \leq \sqrt{m} Lip(f)$

Proof. Assume for now that $m = 1$. Define

$$\overline{f}(x) = \inf\{f(a) + Lip(f) \ |x - a| \ ; \ a \in A\}$$

$$\widetilde{f}(x) = \sup\{f(a) + Lip(f) \ |x - a| \ ; \ a \in A\}$$

Note that $\widetilde{f}$ defined as above is another extension of f, need not be same as $\overline{f}$. Continuing the proof for $\overline{f}$, if $b \in A$, then clearly, $\overline{f}(b) = f(b)$. If $x, y \in \mathbb{R}^n$, then,

$$\overline{f}(x) = \inf\{f(a) + Lip(f)\ |x-a|\ ;\ a \in A\}$$

$$\leq \inf\{f(a) + Lip(f)\ (|y-a| + |y-x|)\ ;\ a \in A\}$$

$$\leq \inf\{f(a) + Lip(f)\ |y-a|\ ;\ a \in A\} + Lip(f)\ |x-y|$$

$$= \overline{f}(y) + Lip(f)\ |x-y|$$

Similarly,

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

Thus

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

In the general case, $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, let $f \equiv (f_1, f_2, \ldots, f_m)$ with $f_i : A \to \mathbb{R}$ for all $1 \leq i \leq m$. By the previous steps, there are $\overline{f}_i$ such that $\overline{f}_i = f$ on A and $Lip(\overline{f}_i) \leq Lip(f)$. Thus

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

$$\implies Lip(\overline{f}) \leq \sqrt{m} Lip(f)$$

$\square$

Theorem 2.63. ***Rademacher's theorem :***
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally lipshitz. Then f is differentiable $\mathcal{L}^n$ almost everywhere.

Proof. Assume that $m = 1$. Also, since the differentiablity is a local property, assume the f is globally lipshitz. Fix $v \in \mathbb{R}^n$ such that $|v| = 1$. Define, provided the limit exist,

$$D_v f(x) := \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}\ ;\ x \in \mathbb{R}^n$$

Claim 1 : $D_v f(x)$ exist for $\mathcal{L}^n$ almost every $x \in \mathbb{R}^n$.
Proof of the claim 1 : Since f is continuous,

$$\overline{D_v} f(x) := \limsup_{t \to 0} \frac{f(x+tv) - f(x)}{t} = \lim_{k \to \infty} \sup_{\{0 < |t| < \frac{1}{k}\ ,\ t \in Q\}} \frac{f(x+tv) - f(x)}{t}$$

is also borel measurable. Similarly,

$$\underline{D_v} f(x) \equiv \liminf_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

is borel measurable. Thus

$$A_v \equiv \{x \in \mathbb{R}^n \ ; \ D_v f(x) \text{ does not exist } \}$$

$$= \{x \in \mathbb{R}^n \ ; \ \underline{D_v} f(x) \neq \overline{D_v} f(x)\}$$

is borel measurable. For each $x, v \in \mathbb{R}^n$ with $|v| = 1$, define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(t) := f(x + tv) \text{ for } t \in \mathbb{R}$$

Then g is lipshitz and thus absolutely continuous and thus differentiable $\mathcal{L}^1$ almost everywhere. Hence,

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

for each line L parallel to v. By the fubini's theorem,

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

This proves the claim 1.
Consequently,

$$\nabla f(x) \equiv \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)(x) \text{ exist for } \mathcal{L}^n \text{ almost every } x$$

Note that sometimes $\nabla f(x)$ is also written as $Df(x)$.
Claim 2 : $D_v f(x) = v.\nabla f(x)$ $\mathcal{L}^n$ a.e x.
Proof of the claim 2 : Let $\xi \in C_c^\infty(\mathbb{R}^n)$. Then,

$$\int_{\mathbb{R}^n} \left(\frac{f(x + tv) - f(x)}{t} \right) \xi(x) dx = - \int_{\mathbb{R}^n} f(x) \left(\frac{\xi(x) - \xi(x - tv}{t} \right) dx$$

Let $t = \frac{1}{k}$, $k = 1, 2, \ldots$. Thus the above equality gives that

$$\left| \frac{f(x + \frac{1}{k}v) - f(x)}{\frac{1}{k}} \right| \leq Lip(f)|v| = Lip(f)$$

Thus, the D.C.T implies,

$$\int_{\mathbb{R}^n} D_v f(x) \xi(x) dx = - \int_{\mathbb{R}^n} f(x) D_v \xi(x) dx$$

$$= - \sum_{i=1}^{\infty} \int_{\mathbb{R}^n} f(x) \frac{\partial \xi}{\partial x_i}(x) \xi(x) dx$$

$$= \sum_{i=1}^{\infty} v_i \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x) \xi(x) dx$$

$$= \int_{\mathbb{R}^n} \big(v.\nabla f(x)\big)\xi(x)dx$$

The above is true as per the theorem of Fubini and the note after corollary 1.72 that says the integration by parts holds for absolutely continuous functions. The above holds for all $\xi \in C_c^\infty(\mathbb{R}^n)$. Hence by the density of the smooth functions with compact support, we have

$$D_v f = v.\nabla f \text{ for } \mathcal{L}^n \text{ a.e}$$

This proves the claim 2.
Let $\{v_k\}_{k=1}^\infty$ be countable dense subset of $\partial B(0,1)$. Define

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

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

Note that, from the claim 1 and 2, $\mathcal{L}^n(\mathbb{R}^n - A) = 0$.
Claim 3 : f is differentiable at each $x \in A$.
Proof of the claim 3 : Choose and fix $x \in A$. Choose $v \in \partial B(0,1)$, $t \in \mathbb{R}$, $t \neq 0$. Define

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

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

$$|Q(x,v,t) - Q(x,u,t)| \leq \left|\frac{f(x+tv) - f(x+tu)}{t}\right| + |(v-u).\nabla f(x)|$$

$$\leq Lip(f)|v-u| + |\nabla f(x)||v-u|$$

$$\leq (\sqrt{n}+1)Lip(f)|v-u|$$

Let $\epsilon > 0$ and $\epsilon_1 := \frac{\epsilon}{2\left(\sqrt{n}+1\right)Lip(f)}$. Then, by the compactness of $\partial B(0,1)$ and dense property of $\{v_k\}$, choose N such that

$$\partial B(0,1) \subset \bigcup_{k=1}^N B(v_k, \epsilon_1)$$

Therefore, for each $v \in \partial B(0,1)$, we can find $k \in \{1,2,\ldots,N\}$ such that $|v - v_k| \leq \epsilon_1$. Since, $\lim_{t\to 0} Q(x, v_k, t) = 0$, $\forall k = 1,2,\ldots,N$. Hence, choose $\delta > 0$ such that $0 < |t| < \delta$, $k = 1,2,\ldots,N$ and

$$|Q(x, v_k, t)| \leq \frac{\epsilon}{2}$$

As a consequence, for each $v \in \partial B(0,1)$, there is a $k \in \{1, 2, \dots N\}$ such that, if $0 < |t| < \delta$,

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

By the compactness argument, the same $\delta > 0$ works for all $v \in \partial B(0,1)$. Now, choose $y \in \mathbb{R}^n$ such that $y \neq x$. Declare

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

$$y = x + tv$$

$$t \equiv |x-y|$$

Then,

$$f(y) - f(x) - \nabla f(x).(y-x) = f(x+tv) - f(x) - tv.\nabla f(x)$$

$$= o(t) \equiv o(|x-y|), \text{ as } y \to x$$

Hence, f is differentiable at x with $Df(x) = \nabla f(x)$. □

Corollary 2.64.

- *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally lipshitz and* *$Z := \{x \in \mathbb{R}^n \ ; \ f(x) = 0\}$. Then $Df(x) = 0$ for $\mathcal{L}^n$ a.e $x \in Z$.*
- *Let $f, g : \mathbb{R}^n \to \mathbb{R}^n$ be locally lipshitz and $Y := \{x \in \mathbb{R}^n \ ; \ g \circ f(x) = x\}$. Then*

$$Dg(f(x)) \ Df(x) = I \text{ for } \mathcal{L}^n \text{ a.e } x \in Y$$

Proof. Assume that $m = 1$. Let $x \in Z$ be a point of density and $Df(x)$ exist, that is,

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

Then,

$$f(y) = Df(x).(y-x) + o(|y-x|) \text{ as } y \to x$$

Now, assume that $Df(x) \equiv b \neq 0$ and declare

$$S := \{v \in \partial B(0,1) \ ; \ b.v > \frac{1}{2}|b|\}$$

$$V(r,b) := \{tv \ ; \ 0 < t < r \ , \ v \in S\}$$

Then S is open neighbourhood of $\frac{b}{|b|}$ and $V(r,b)$ is an open set with $V(r,b) \subset B(x,r)$ and

$$\alpha = \frac{\mathcal{L}^n\big(V(1,b)\big)}{\mathcal{L}^n\big(B(x,1)\big)} = \frac{r^n\mathcal{L}^n\big(V(1,b)\big)}{r^n\mathcal{L}^n\big(B(x,1)\big)} = \frac{\mathcal{L}^n\big(V(r,b)\big)}{\mathcal{L}^n\big(B(x,r)\big)} < 1$$

For each $v \in S$ and $t > 0$, set $y = x + tv$ to get

$$f(x+tv) = b.tv + o(|tv|) \geq \frac{t|b|}{2} + o(t) \text{ as } t \to 0$$

Hence, there is $t_0 > 0$ such that

$$f(x+tv) > 0 \text{ , for } 0 < t < t_0 \text{ , } v \in S.$$

Hence, for $0 < r < t_0$, $V(r,b) \cap Z = \phi$.

$$\implies \frac{\mathcal{L}^n\big(Z \cap B(x,r)\big)}{\mathcal{L}^n\big(B(x,r)\big)} = \frac{\mathcal{L}^n\Big(Z \cap \big(B(x,r) - V(r,b)\big)\Big)}{\mathcal{L}^n\big(B(x,r)\big)} \leq \frac{\mathcal{L}^n\big(B(x,r)\big) - \mathcal{L}^n\big(V(r,b)\big)}{\mathcal{L}^n\big(B(x,r)\big)}$$

$$= 1 - \alpha < 1$$

Letting $r \to 0$, to get a contradiction, as x is a point of density.

To prove the second part of the corollary, declare

$$S(Df) := \{x \text{ ; } Df(x) \ exists\}$$

$$S(Dg) := \{x \text{ ; } Dg(x) \ exists\}$$

$$X := Y \cap S(Df) \cap f^{-1}\Big(S\big(Dg\big)\Big)$$

Then,

$$x \in Y - f^{-1}(S(Dg)) \implies f(x) \in \mathbb{R}^n - S(Dg) \implies x = g \circ f(x) \in g(\mathbb{R}^n - S(Dg)$$

$$\implies Y - X \subset \Big(\mathbb{R}^n - S(Df)\Big) \cup g\Big(\mathbb{R}^n - S(Dg)\Big)$$

By the Rademacher's theorem,

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

Thus, if $x \in X$, then $Df(x)$ exists , $Dg(f(x))$ exists and so,

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

Observing that $g \circ f \equiv I$ on Y, and the previous part of this corollary tells that

$$D(g \circ f) \equiv I \text{ , } \mathcal{L}^n \text{ a.e on } Y.$$

$\square$

2.4 Area formula

2.4.1 Theorem and proofs.

Throughout this section (2.4) , assume that $n \leq m$.

Lemma 2.65. *Suppose that $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then*

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

Proof. Write $L = O \circ S$.
Case 1 : If $[[L]] = 0$. Then $dim(S(\mathbb{R}^n)) \leq n-1$ and hence, $dim(L(\mathbb{R}^n)) \leq n-1$. Thus as a consequence, $\mathcal{H}^n(L(\mathbb{R}^n)) = 0$.
Case 2 : Let $[[L]] > 0$, , then by the change of variables formula and the fact from the isodiametric inequality, proved earlier that $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$,

$$\frac{\mathcal{H}^n\big(L(B(x,r))\big)}{\mathcal{L}^n\big(B(x,r)\big)} = \frac{\mathcal{H}^n\Big(O^* \circ L\big(B(x,r)\big)\Big)}{\mathcal{L}^n\big(B(x,r)\big)}$$

$$= \frac{\mathcal{H}^n\Big(O^* \circ O \circ S\big(B(x,r)\big)\Big)}{\mathcal{L}^n\big(B(x,r)\big)}$$

$$= \frac{\mathcal{H}^n\Big(S\big((B(x,r)\big)\Big)}{\mathcal{L}^n\big(B(x,r)\big)}$$

$$= \frac{\mathcal{L}^n\Big(S\big((B(x,r)\big)\Big)}{\mathcal{L}^n\big(B(x,r)\big)}$$

$$= \frac{\mathcal{L}^n(S(B(0,1)))}{\alpha(n)}$$

$$= |detS|$$

$$= [[L]]$$

Define $v(A) := \mathcal{H}^n(L(A))$ for $A \subset \mathbb{R}^n$. Then clearly,

$$v \text{ is a radon measure}$$

$$v \ll \mathcal{L}^n$$

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

Hence, for all borel sets $B \subset \mathbb{R}^n$, by the Radon- Nikodym theorem 2.14,

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

Noting that v and $\mathcal{L}^n$ are radon measures, the same formula holds for all sets $A \subset \mathbb{R}^n$. $\square$

Henceforth, for this section, it shall be assumed that f is lipshitz mapping.

Lemma 2.66. *Let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$ measurable. Then*

- *$f(A)$ is $\mathcal{H}^n$ measurable*
- *Multiplicity function : $y \to \mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$ measurable on $\mathbb{R}^m$*
- *$\int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\}) d\mathcal{H}^n \leq (Lip(f))^n \mathcal{L}^n(A)$*

Proof. W.L.O.G, assume that A is bounded,i.e $\mathcal{L}^n(A) < \infty$. By the approximation of sets by compact sets from the inside, there are $K_i \subset A$ such that, for all $i \in \mathbb{N}$,

$$\mathcal{L}^n(A - K_i) = \mathcal{L}^n(A) - \mathcal{L}^n(K_i) \leq \frac{1}{i}$$

$$\implies \mathcal{L}^n(K_i) \geq \mathcal{L}^n(A) - \frac{1}{i}$$

Since, f is continuous, $f(K_i)$ is compact and thus is $\mathcal{H}^n$ measurable. Thus

$$f\left(\bigcup_{i=1}^{\infty} K_i\right) = \bigcup_{i=1}^{\infty} f(K_i) \text{ is } \mathcal{H}^n \text{ measurable.}$$

Furthermore,

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

$$\leq (Lip(f))^n \mathcal{L}^n\left(A - \bigcup_{i=1}^{\infty} K_i\right) = 0$$

thus $f(A)$ is $\mathcal{H}^n$ measurable and this proves the first part of the lemma.
Declare

$$B_k := \{Q \mid Q \equiv (a_1, b_1] \times \ldots \times (a_n, b_n] \ , \ a_i = \frac{c_i}{k} \ , \ b_i = \frac{c_i + 1}{k} \ , c_i \in \mathbb{Z} \ , \ i \in \{1, 2, \ldots, n\}\}$$

Observe that $\mathbb{R}^n = \bigcup_{Q \in B_k} Q$. Now, define $g_k := \sum_{Q \in B_k} \mathcal{X}_{f(A \cap Q)}$. By the first part of this lemma, g_k is $\mathcal{H}^n$ measurable. Observe that $g_k(y)$ is the number of cubes $Q \in B_k$

such that $f^{-1}(y) \cap (A \cap Q) \neq \phi$. Thus $g_k(y) \uparrow \mathcal{H}^0(A \cap f^{-1}(y))$ as $k \to \infty$, for each $y \in \mathbb{R}^m$. This proves the second point of this lemma.
By MCT,

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

□

Lemma 2.67. *Let*

$$t > 1$$

$$B := \{x \; ; \; Df(x) \; exists \; , \; Jf(x) > 0\}$$

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

- $B = \bigcup_{k=1}^{\infty} E_k$
- f *is injective on* E_k.
- *for each* $k \in \mathbb{N}$, *there are symmetric automorphisms* $T_k : \mathbb{R}^n \to \mathbb{R}^n$ *such that*

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

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

$$\frac{1}{t^n} |detT_k| \leq Jf|_{E_k} \leq t^n |detT_k|$$

Proof. Fix $\epsilon > 0$ such that $\frac{1}{t} + \epsilon < 1 < t - \epsilon$. Let $\mathcal{S}$ be a countable dense subset of B and let $\mathcal{T}$ be a countable dense set of symmetric automorphisms of $\mathbb{R}^n$. Then for each $s \in \mathcal{S}, T \in \mathcal{T}$, define for $i \in \mathbb{N}$,

$$E(s,T,i) := \left\{ b \in B \cap B(s, \tfrac{1}{i}) \;\middle|\; \begin{array}{c} (1) \left(\frac{1}{t} + \epsilon\right) |Tv| \leq |Df(b)v| \leq (t - \epsilon)|Tv|, \forall v \in \mathbb{R}^n \; , \\ (2) \; |f(a) - f(b) - Df(b).(a - b)| \leq \epsilon |T(a - b)|, \forall a \in B(b, \frac{2}{i}) \end{array} \right\}$$

Note that $b \to Df(b)$ and $v \to Df(b)v$ are Borel measurable as f is lipshitz. Further calculations of $(1), (2)$ lead to the fact that

$$\frac{1}{t}|T(a - b)| \leq |f(a) - f(b)| \leq t|T(a - b)| \; , \text{for all } b \in E(s, T, i), a \in B(b, \frac{2}{i})$$

Claim :

$$b \in E(s,T,i) \implies \left(\frac{1}{t}+\epsilon\right)^n |detT| \le Jf(b) \le (t-\epsilon)^n |detT|$$

Proof of the claim : Declare $Df(b) = L = O \circ S$ Then,

$$Jf(b) = [[Df(b)]] = |detS|$$

By (1),

$$\left(\frac{1}{t}+\epsilon\right)|Tv| \le |O \circ S(v)| = |Sv| \le (t-\epsilon)|Tv| \text{ , for } v \in \mathbb{R}^n$$

$$\implies \left(\frac{1}{t}+\epsilon\right)|v| \le |S \circ T^{-1}(v)| \le (t-\epsilon)|v| \text{ , for } v \in \mathbb{R}^n$$

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

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

The other inequality is similar in proof. This proves the claim,

Relabel $E(s,T,i)$ as $\{E_k\}_{k\in\mathbb{N}}$. Select any $b \in B$ and write $Df(b) = O \circ S$ as above and choose $T \in \mathcal{T}$ such that

$$Lip\big(T \circ S^{-1}\big) \le \left(\frac{1}{t}+\epsilon\right)^{-1}$$

$$Lip\big(S \circ T^{-1}\big) \le t-\epsilon$$

Now select $i \in \mathbb{N}$ and $c \in \mathcal{S}$ so that

$$|b-c| \le \frac{1}{i}$$

$$|f(a) - f(b) - Df(b).(a-b)| \le \frac{\epsilon}{Lip\,(T^{-1})}|a-b| \le \epsilon|T(a-b)|\text{, for all } a \in B(b, \frac{2}{i})$$

Then, $b \in E(c,T,i)$. This proves the first part of the lemma, i.e , $B = \bigcup E_k$.
Now, choose any $E_k \equiv E(c,T,i)$. Thus by the observation,

$$\frac{1}{t}|T(a-b)| \le |f(a)-f(b)| \le t.|T(a-b)| \text{ , for all } b \in E_k, a \in B(b, \frac{2}{i})$$

As, $E_k \subset B(b, \frac{2}{i})$, we have

$$\frac{1}{t}|T(a-b)| \le |f(a)-f(b)| \le t.|T(a-b)| \text{ , for all } a,b \in E_k$$

This tells the second point of the lemma,i.e f is injective on each E_k.
Finally the above inequality gives

$$Lip\Big((f|_{E_k}) \circ T^{-1}\Big) \leq t$$

$$Lip\Big(T \circ (f|_{E_k})^{-1}\Big) \leq t$$

Now, for all $a \in B(b, \frac{2}{i})$, we have

$$|f(a) - f(b)| \leq t|T(a-b)|$$

let $v \in \mathbb{R}^n$, non zero, and $a = b + sv$ such that $|v||s| < \frac{2}{i}$, then

$$|f(b+sv) - f(b)| \leq t|s||T(v)|$$

and hence,

$$\left|\frac{f(b+sv) - f(b)}{s}\right| \leq t|Tv|$$

Letting $s \to 0$,

$$\implies |Df(b)v| \leq t|Tv|$$

$$|S \circ T^{-1}(v)| = |O \circ S \circ T^{-1}(v)| = |Df(b) \circ T^{-1}(v)| \leq t|v|$$

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

$$\implies |detS||detT^{-1}| \leq t^n$$

That is

$$|Jf(b)| \leq t^n |detT|$$

Similarly, considering the other inequality , we get

$$|Jf(b)| \geq \frac{1}{t^n}|detT|$$

This proves the lemma.

□

Theorem 2.68. ***Area formula :***
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be lipshitz ,$n \leq m$, then for each $\mathcal{L}^n$ measurable subset $A \subset \mathbb{R}^n$,

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

Proof. In the view of the Rademacher's theorem, we may assume that $Df(x)$ and $Jf(x)$ exists on whole of A. Also, assume that $\mathcal{L}^n(A) < \infty$.
Case 1 :

$$A \subset \{Jf > 0\}$$

Choose and fix $t > 1$. Now choose borel sets E_k as in the previous lemma. Make them disjoint by the usual strategy of declaring $F_1 = E_1$ and for all $i > 1$, declare $F_i = E_i - \bigcup_{j=1}^{i-1} E_j$.
Define, as in one of the earlier lemmas,

$$B_k := \{Q \mid Q \equiv (a_1, b_1] \times \ldots \times (a_n, b_n] \ , \ a_i = \frac{c_i}{k} \ , \ b_i = \frac{c_i + 1}{k} \ , c_i \in \mathbb{Z} \ , \ i \in \{1, 2, \ldots, n\}\}$$

Now, set

$$F_{ij} = F_j \cap Q_i \cap A \ , \text{ for } Q_i \in B_k, j \in \mathbb{N}$$

By the construction, F_{ij}'s are disjoint. Also,

$$A = \bigcup_{i,j=1}^{\infty} F_{ij}$$

Claim 1:

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

Proof of the claim 1: Declare

$$g_k := \sum_{i,j=1}^{\infty} \mathcal{X}_{f(F_{ij})}$$

Observe that $g_k(y)$ is the exact number of the sets F_{ij} such that $F_{ij} \cap f^{-1}\{y\} \neq \phi$. As the cubes gets finer as $k \to \infty$, by the MCT,

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

This concludes the claim.
Now, by the previous lemma,

$$\mathcal{H}^n\Big(f(F_{ij})\Big) = \mathcal{H}^n\Big(f|_{E_j} \circ T_j^{-1} \circ T_j(F_{ij})\Big) \leq t^n \mathcal{L}^n\Big(T_j(F_{ij})\Big)$$

$$\mathcal{L}^n\Big(T_j(F_{ij})\Big) = \mathcal{H}^n\Big(T_j \circ \big(f|_{E_j}\big)^{-1} \circ f\big(F_{ij}\big)\Big) \leq t^n \mathcal{H}^n\Big(f(F_{ij})\Big)$$

Thus

$$\frac{1}{t^{2n}} \mathcal{H}^n\Big(f(F_{ij})\Big) \leq \frac{1}{t^n} \mathcal{L}^n\Big(T_j(F_{ij})\Big)$$

$$= \frac{1}{t^n} |detT_j| \mathcal{L}^n(F_{ij})$$

$$\leq \int_{F_{ij}} Jf dx$$

$$\leq t^n |detT_j| \mathcal{L}^n(F_{ij})$$

$$= t^n \mathcal{L}^n \Big(T_j(F_{ij}) \Big)$$

$$\leq t^{2n} \mathcal{H}^n \Big(f(F_{ij}) \Big)$$

Summing the above over i, j,

$$\frac{1}{t^{2n}} \sum_{i,j=1}^{\infty} \mathcal{H}^n \Big(f(F_{ij}) \Big) \leq \int_A Jf dx \leq t^{2n} \sum_{i,j=1}^{\infty} \mathcal{H}^n \Big(f(F_{ij}) \Big)$$

By the claim 1, as $k \to \infty$,

$$\frac{1}{t^{2n}} \int_{\mathbb{R}^m} \mathcal{H}^0 \Big(A \cap f^{-1}\{y\} \Big) d\mathcal{H}^n \leq \int_A Jf dx \leq t^{2n} \int_{\mathbb{R}^m} \mathcal{H}^0 \Big(A \cap f^{-1}\{y\} \Big) d\mathcal{H}^n$$

Sending $t \to 1+$, proves the theorem for the Case 1.

Case 2 :

$$A \subset \{Jf = 0\}$$

Choose and fix $0 < \epsilon \leq 1$. Write f as $f = p \circ g$, where

$$g : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n \ ; \ x \to \big(f(x), \epsilon.x\big)$$

$$p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \ ; \ (y, z) \to y$$

Claim 2 : There is a constant C such that

$$0 < Jg(x) \leq C\epsilon \text{ for } x \in A$$

Proof of the claim 2 : Write $g = (f_1, f_2, \ldots, f_m, \epsilon x_1, \epsilon x_2, \ldots, \epsilon x_n)$, then

$$Dg(x) \equiv \begin{bmatrix} Df(x) \\ \epsilon I \end{bmatrix}_{(n+m)\times n}$$

Note that, by the Binet-Cauchy formula, the $Jf^2(x)$ is equal to the sum of the determinant of all the $n \times n$ sub-matrices of $Df(x)$. Thus $Jg^2(x) \geq \epsilon^{2n} > 0$. Furthermore, since, $|Df| \leq Lip(f) < \infty$, by the Binet-Cauchy formula, $Jg^2(x) = Jf^2(x) + \{$ sum of squares of terms involving atleast one $\epsilon\} \leq C.\epsilon^2$, for each $x \in A$. Since p is a projection, by the case 1 ,

$$\mathcal{H}^n\Big(f(A)\Big) \le \mathcal{H}^n\Big(g(A)\Big)$$

$$\le \int_{\mathbb{R}^{n+m}} \mathcal{H}^0\Big(A \cap g^{-1}\{y,z\}\Big) d\mathcal{H}^n(y,z)$$

$$= \int_A Jg(x)dx$$

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

Let $\epsilon \to 0$ to conclude that $\mathcal{H}^n\Big(f(A)\Big) = 0$. Thus, as *support*$\Big(\mathcal{H}^0\big(A \cap f^{-1}\{y\}\big)\Big) \subset f(A)$,

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

In the general case, split any set A as union of $A_k = A \cap \Big(B(0,k) - B(0,k-1)\Big)$, $A_1 = A \cap B(0,1)$ and split A_k as $A_{k,1} \bigcup A_{k,2}$ where

$$A_{k,1} \subset \{Jf > 0\}$$

$$A_{k,2} \subset \{Jf = 0\}$$

Apply the case (1) and (2), to conclude the theorem. □

Theorem 2.69. *Change of Variables :*
Let $f : \mathbb{R}^n \to \mathbb{R}^m$, be lipshitz, $n \le m$, then, for each $\mathcal{L}^n$ integrable function $g : \mathbb{R}^n \to \mathbb{R}$,

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

Proof. Idea is to prove it for simple functions and then use the limit. Consider $g = \sum_{finite} C_i \mathcal{X}_{A_i}$. Thus

$$\int_{\mathbb{R}^n} g(x)Jf(x)dx = \sum \int_{\mathbb{R}^n} C_i \mathcal{X}_{A_i} Jf(x)dx$$

$$= \sum C_i \int_{A_i} Jf(x)dx$$

$$= \sum C_i \int_{\mathbb{R}^m} \mathcal{H}^0\Big(A_i \cap f^{-1}\{y\}\Big) d\mathcal{H}^n(y)$$

$$= \int_{\mathbb{R}^m} \sum_i C_i \sum_{x \in f^{-1}\{y\}} \mathcal{X}_{A_i}(x) d\mathcal{H}^n(y)$$

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

Now, g is integrable implies it can be approximated by simple functions which are integrable and hence, by DCT, the theorem is proved. □

2.4.2 Applications:

- Length of a Curve.

  Here $n = 1, m \geq 1$. Assume that $f : \mathbb{R} \to \mathbb{R}^m$ is lipshitz and one-one.

  Note that if $f = (f_1, f_2, \ldots, f_m)$, then

$$Df = (f_1', f_2', \ldots f_m')$$

$$Jf = |Df| = |f'| = \sqrt{\sum_{i=1}^{m} |f_i'|^2}$$

  For $-\infty < a < b < \infty$, define the curve $S \equiv f\big([a, b]\big) \subset \mathbb{R}^m$. Note that the summand in the area formula is just one term due to the injectiveness of the f. Then the Length of the curve S is

$$\mathcal{H}^1(S) = \int_a^b |f'(t)| dt$$

- Surface area of a graph :

  Here , $n \geq 1, m = n + 1$. Assume $g : \mathbb{R}^n \to \mathbb{R}$ is lipshitz and define, $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

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

  Then we have the following

$$Df = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \frac{\partial g}{\partial x_3} & \ldots & \frac{\partial g}{\partial x_n} \end{bmatrix}_{(n+1)\times n}$$

  And

$$(Jf)^2 = 1 + |Dg|^2$$

  For each open subset of $\mathbb{R}^n$, define the graph of g over U

$$G \equiv G(g, U) = \{(x, g(x) \ ; \ x \in U\} \subset \mathbb{R}^{n+1}$$

Then, the surface area of the graph of g, that is surface area of G is

$$\mathcal{H}^n(G) = \int_U \sqrt{1 + |Dg|^2} dx$$

- Surface area of a parametric hypersurface.
  Here $n \geq 1, m = n + 1$. Assume that $f : \mathbb{R}^n \to \mathbb{R}^{n+1}$ is lipshitz and one-one. Write $f = (f_1, f_2, \ldots, f_{n+1})$, then

$$Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n+1}}{\partial x_1} & \cdots & \frac{\partial f_{n+1}}{\partial x_n} \end{bmatrix}$$

Let

$$T_k := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{k-1}}{\partial x_1} & \cdots & \frac{\partial f_{k-1}}{\partial x_n} \\ \frac{\partial f_{k+1}}{\partial x_1} & \cdots & \frac{\partial f_{k+1}}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{n+1}}{\partial x_1} & \cdots & \frac{\partial f_{n+1}}{\partial x_n} \end{bmatrix}$$

$$\implies (Jf)^2 = \sum_{k=1}^{n+1} det(T_k)^2$$

For each U, open subset of $\mathbb{R}^n$, let the hypersurface be $A = f(U) \subset \mathbb{R}^{n+1}$. Then, the surface area of A is,

$$\mathcal{H}^n(A) = \int_U \sqrt{\sum_{k=1}^{n+1} \Big(det(T_k)^2 \Big)}$$

- Submanifolds:
  Let $M \subset \mathbb{R}^m$ be a lipshitz , $n-$ dimensional embedded submanifold. Suppose that $U \subset \mathbb{R}^n$ and $f : U \to M$ is a chart . Let $A \subset f(U)$, A is borel, $B = f^{-1}(A)$, define,

$$g_{ij} = \frac{\partial f}{\partial x_i} . \frac{\partial f}{\partial x_j}$$

$$g = det\big(g_{ij}\big)$$

Then,

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

$$Jf = \sqrt{g}$$

Thus the volume of A in submanifold M is

$$\mathcal{H}^n(A) = \int_B \sqrt{g}dx$$

2.5 Co-Area Formula

2.5.1 Theorems and Proofs.

Throughout this section (2.5), it is assumed that $n \geq m$

and

$f : \mathbb{R}^n \to \mathbb{R}^m$ is a lipshitz mapping.

Lemma 2.70. *Suppose that $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, $n \geq m$ and $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$ measurable, then*

- $y \to \mathcal{H}^{n-m}\big(A \cap L^{-1}\{y\}\big)$ *is $\mathcal{L}^m$ measurable*
- $\int_{\mathbb{R}^m} \mathcal{H}^{n-m}\big(A \cap L^{-1}\{y\}\big)dy = [[L]]\mathcal{L}^n(A)$

Proof. Case 1 :

$$dim(L(\mathbb{R}^n) < m$$

Then,

$$A \cap L^{-1}\{y\} = \phi \implies \mathcal{H}^{n-m}\big(A \cap L^{-1}\{y\}\big) = 0 \text{ , for } \mathcal{L}^m \text{ a.e } y \in \mathbb{R}^m$$

If, L is decomposed as $S \circ O^*$ as in the polar decomposition, then

$$L(\mathbb{R}^n) = S(\mathbb{R}^m)$$

Thus $dim\big(S(\mathbb{R}^m)\big) < m$ and thus

$$[[L]] = |detS| = 0$$

Case 2:

$$L = P = \text{ orthogonal projection of } \mathbb{R}^n \text{ onto } \mathbb{R}^m$$

Then for each $y \in \mathbb{R}^m$, $P^{-1}\{y\}$ is some translated $n-m$ dimensional subspace of $\mathbb{R}^n-$ a translate of $P^{-1}\{0\}$. Thus, by the fubini's theorem,

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}\big(A \cap P^{-1}\{y\}\big)dy = \mathcal{L}^n(A)$$

and thus

$$y \to \mathcal{H}^{n-m}\big(A \cap P^{-1}\{y\}\big) \text{ is } \mathcal{L}^m \text{ measurable.}$$

Case 3 :

$$L : \mathbb{R}^n \to \mathbb{R}^m \ , \ dim\big(L(\mathbb{R}^n)\big) = m$$

By the polar decomposition, $L = S \circ O^*$ such that $[[L]] = |detS| > 0$.
Claim :

$$O^* = P \circ Q$$

where

$$P : \mathbb{R}^n \to_{onto} \mathbb{R}^m \ , \text{ orthogonal projection}$$

$$Q : \mathbb{R}^n \to \mathbb{R}^n \ , \text{ orthogonal}$$

Proof of the claim : Choose $Q : \mathbb{R}^n \to \mathbb{R}^n$, orthogonal such that

$$Q^*(x) = O(x) \ , \text{ for all } x \in \mathbb{R}^m$$

Noting that

$$P^*(x_1, x_2, \ldots, x_m) = (x_1, x_2, \ldots, x_m, 0, 0 \ldots, 0) \in \mathbb{R}^n \ , \text{ for all } x \in \mathbb{R}^m$$

Thus, we have

$$O = Q^* \circ P^*$$

$$\implies O^* = P \circ Q$$

Observe that $L^{-1}\{0\}$ is a $n-m$ dimensional subspace of $\mathbb{R}^n$ and as before, $L^{-1}\{y\}$ is a translate of $L^{-1}\{0\}$. By the fubini's theorem,

$$\mathcal{L}^n(A) = \mathcal{L}^n\big(Q(A)\big) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(Q(A) \cap P^{-1}\{y\}\Big)dy$$

$$= \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A \cap Q^{-1} \circ P^{-1}\{y\}\Big)dy$$

By the area formula's change of variables , let $z = Sy$, then

$$|detS|\mathcal{L}^n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A \cap Q^{-1} \circ P^{-1} \circ S^{-1}\{y\}\Big)dy$$

Noting that $L = S \circ O^* = S \circ P \circ Q$,

$$\implies [[L]]\mathcal{L}^n(A) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A \cap L^{-1}\{z\}\Big)dz$$

Thus

$$y \to \mathcal{H}^{n-m}\Big(A \cap L^{-1}\{y\}\Big) \text{ is } \mathcal{L}^m \text{ measurable.}$$

□

Lemma 2.71. *Let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$- measurable, $n \geq m$. Then,*

- *$A \cap f^{-1}\{y\}$ is $\mathcal{H}^{n-m}$ measurable, for $\mathcal{L}^m$ almost every y in $\mathbb{R}^m$.*
- *$y \to \mathcal{H}^{n-m}\big(A \cap f^{-1}\{y\}\big)$ is $\mathcal{L}^m$ measurable*
- $$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A \cap f^{-1}\{y\}\Big)dy \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}\Big(Lip(f)\Big)^m \mathcal{L}^n(A)$$

Proof. Assume A is bounded. Else, do it for $A \cap B(0,k)$, $k \in \mathbb{N}$. For each $j \in \mathbb{N}$,by the definition of $\mathcal{L}^n$, there are closed balls $\{B_{ij}\}_{i=1}^{\infty}$ such that

- $A \subset \bigcup_{i=1}^{\infty} B_{ij}$
- $diam\big(B_{ij}\big) \leq \frac{1}{j}$
- $\sum_{i=1}^{\infty} \mathcal{L}^n\big(B_{ij}\big) \leq \mathcal{L}^n(A) + \frac{1}{j}$

Define g_{ij} as

$$g_{ij}(x) := \alpha(n-m)\bigg(\frac{diam(B_{ij}}{2}\bigg)^{n-m} \mathcal{X}_{f(B_{ij})}(x)$$

By the first part of the lemma, g_{ij} is measurable. Now, observe that for all $y \in \mathbb{R}^m$,

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

Note, that it is not yet clear that $y \to \mathcal{H}^{n-m}\big(A \cap f^{-1}((y)\big)$ is measurable, as of now, in this proof. Denote the upper lebesgue integral as

$$\int^*_{\mathbb{R}^m} h(y)dy := \inf\{\int_{\mathbb{R}^m} \Phi(y)dy \ ; \ \phi \text{ is simple } , \ h(y) \leq \Phi(y)\}$$

Thus,

$$\int^*_{\mathbb{R}^m} \mathcal{H}^{n-m}\big(A \cap f^{-1}\{y\}\big)dy$$

$$= \int^*_{\mathbb{R}^m} \lim_{j\to\infty} \mathcal{H}^{n-m}_{\frac{1}{j}}\big(A \cap f^{-1}\{y\}\big)dy$$

$$\leq \int^*_{\mathbb{R}^m} \lim_{j\to\infty} \sum_{i=1}^{\infty} g_{ij}dy$$

By the fatou's lemma and since g_{ij} is measurable,

$$\leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \int_{\mathbb{R}^m} g_{ij} dy$$

$$= \liminf_{j\to\infty} \sum_{i=1}^{\infty} \alpha(n-m)\left(\frac{diam(B_{ij})}{2}\right)^{n-m} \mathcal{L}^m\big(f(B_{ij})\big)$$

By the isodiamteric property,

$$\leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \alpha(n-m)\left(\frac{diam(B_{ij})}{2}\right)^{n-m} \alpha(m)\left(\frac{diam\big(f(B_{ij})\big)}{2}\right)^{m}$$

$$= \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}\big(Lip(f)\big)^m \liminf_{j\to\infty} \sum_{i=1}^{\infty} \mathcal{L}^n(B_{ij})$$

$$\leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}\big(Lip(f)\big)^m \mathcal{L}^n(A)$$

And thus

$$\int_{\mathbb{R}^m}^{*} \mathcal{H}^{n-m}\big(A\cap f^{-1}\{y\}\big)dy \leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)}\big(Lip(f)\big)^m \mathcal{L}^n(A) \qquad (*)$$

Thus the idea is the prove the first point. Once the first point is proved, then Fubini and Fatou's lemma implies the second and the third point as per the previous calculations! Now to show the first point of the theorem.
Case 1 :

A is compact.

Choose and fix $t \geq 0$. For each $i \in \mathbb{N}$, let U_i consist of all $y \in \mathbb{R}^m$ such that, there are finitely many open sets S_i, say $i = 1, 2, \ldots, l$ with the property that

- $A \cap f^{-1}\{y\} \subset \bigcup_{i=1}^{l} S_i$
- $diam(S_k) \leq \frac{1}{i}$, for all $k = 1, 2, \ldots, l$
- $\sum_{k=1}^{l} \alpha(n-m)\left(\frac{diam(S_k)}{2}\right)^{n-m} < t + \frac{1}{i}$

Claim 1 :

U_i is open.

Proof of the claim 1 : It is equivalent to show the following

$W = \{y\ ;\ A \cap f^{-1}(y) \subset U\}$ is open with f to be continuous and A is compact.

Let $y_0 \in W$. It is required to show the existence of $\epsilon > 0$ such that $B(y_0, \epsilon) \subset W$. Suppose not, then there is sequences $\{y_k\}, \{\epsilon_k\}$ such that

- $y_k \in B(y_0, \epsilon_k)$
- $y_k \notin W$
- $\epsilon_k \to 0$ as $k \to \infty$
- $y_k \to y_0$

Thus,

$$A \cap f^{-1}(y_k) \not\subset U \implies \exists x_k \in A \cap f^{-1}(y_k)\ , x_k \notin U\ ,\ x_k \in A$$

Note that $x_k \to x_0$ in A, because of the compactness. Also, observe that $f(x_k) = y_k$ and continuity of f implies $f(x_0) = y_0$ with $x_0 \in U^c$.

$$\implies x_o \in A \cap f^{-1}(y_0) \subset U$$

This is a contradiction and hence the claim 1 is true.
Claim 2 :

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

Proof of the claim 2 : By the definiton of supremum of the $\mathcal{H}^{n-m}$, if $\mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) \leq t$, then for all $\delta > 0$, $\mathcal{H}^{n-m}_{\delta}\big(A \cap f^{-1}(y)\big) \leq t$. So, given $i \in \mathbb{N}$, choose $\delta \in (0, \frac{1}{i})$. Then, by the definition of $\mathcal{H}^{n-m}$, there are sets $\{S_j\}_{j=1}^{\infty}$ such that

- $A \cap f^{-1}(y) \subset \bigcup_{j=1}^{\infty} S_j$
- $diam S_j < \delta < 1/i$
- $\sum_{j=1}^{\infty} \alpha(n-m) \left(\frac{diam S_j}{2}\right)^{n-m} < t + \frac{1}{i}$

Since $diam(S_j) = diam(\overline{S_j})$, we can consider S_j's to be closed and let $\eta > 0$, $S_{j\eta} := \{x \in \mathbb{R}^n\ ;\ d(x, S_j) < \eta\}$. Then, $S_j \subset S_{j\eta}$, $S_{j\eta}$ is open and $diam(S_{j\eta}) \to diam(S_j)$ as $\eta \to 0$. Hence, by choosing η small and replacing S_j by $S_{j\eta}$, we can assume that S_j's are open.
Now,since, $A \cap f^{-1}(y)$ is compact, there are finite subcollection, $\{S_i\}_{i=1}^{l}$ such that it covers $A \cap f^{-1}(y)$ and hence, $y \in U_i$. This proves one way inequality, that is

$$\{y\ ;\mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) \leq t\} \subset \bigcap_{i=1}^{\infty} U_i$$

For the other side inequality, if $y \in \cap_{i=1}^{\infty} U_i$, then, for each $i \in \mathbb{N}$,

$$\mathcal{H}^{n-m}_{\frac{1}{i}}\big(A \cap f^{-1}(y)\big) \leq t + \frac{1}{i}$$

$$\Longrightarrow \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) \leq t$$

$$\Longrightarrow \bigcap_{i=1}^{\infty} U_i \subset \{y \ ; \ \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) \leq t\}$$

This proves the claim 2.

The claim 2 also tells that the mapping

$$y \to \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) \text{ is a borel function.}$$

Case 2 :

$$A \text{ is open.}$$

By the inner approximation by the compact sets, there are compact sets $K_1 \subset K_2 \subset K_3 \subset \cdots \subset A$ such that

$$A = \bigcup_{i=1}^{\infty} K_i$$

$$\Longrightarrow \forall y \in \mathbb{R}^m \ , \ \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) = \lim_{i\to\infty} \mathcal{H}^{n-m}\big(K_i \cap f^{-1}(y)\big)$$

$$\Longrightarrow \Big(y \to \mathcal{H}^{n-m}(A \cap f^{-1}(y))\Big) \text{ is a borel function.}$$

Case 3 :

$$\mathcal{L}^n(A) < \infty$$

By the outer regularity of $\mathcal{L}^n$, there are open sets $V_1 \supset V_2 \supset V_3 \supset \cdots \supset A$ such that

$$\mathcal{L}^n(V_1) < \infty \quad ; \quad \mathcal{L}^n\big(\cap_{i=1}^{\infty} V_i - A\big) = 0$$

Note that $\mathcal{L}^n(V_i - A) \to 0$ as $i \to \infty$. Hence,

$$\mathcal{H}^{n-m}\big(V_i \cap f^{-1}(y)\big) \leq \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big) + \mathcal{H}^{n-m}\big((V_i - A) \cap f^{-1}(y)\big)$$

$$\Longrightarrow \limsup_{i\to\infty} \int_{\mathbb{R}^m}^{*} \big|\mathcal{H}^{n-m}\big(V_i \cap f^{-1}(y)\big) - \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big)\big| dy$$

$$\leq \limsup_{i\to\infty} \int_{\mathbb{R}^m}^{*} \mathcal{H}^{n-m}\Big((V_i - A) \cap f^{-1}(y)\Big) dy$$

By (*),

$$\leq \limsup_{i\to\infty} \frac{\alpha(n-m)\alpha(m)}{\alpha(n)} \big(Lip(f)\big)^m \mathcal{L}^n(V_i - A) = 0$$

As a consequence,

$$\mathcal{H}^{n-m}\Big(V_i \cap f^{-1}(y)\Big) \to \mathcal{H}^{n-m}\Big(A \cap f^{-1}(y)\Big)$$

Since, $y \to \mathcal{H}^{n-m}\big(V_i \cap f^{-1}(y)\big)$is $\mathcal{L}^m$ measurable and hence,$y \to \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big)$ is $\mathcal{L}^m$ measurable.
Claim 3 : $A \cap f^{-1}(y)$ is $\mathcal{H}^{n-m}$ measurable.
Proof of the claim 3 : Since $\mathcal{L}^n(V_i - A) \to 0$ as $i \to \infty$,

$$\implies \mathcal{H}^{n-m}\Big((V_i - A) \cap f^{-1}(y)\Big) \to 0 \ ; \ as \ i \to \infty \ , \text{ for } \mathcal{L}^m \text{ a.e } y \in \mathbb{R}^m$$

Let

$$W := \bigcap_{i=1}^{\infty} V_i$$

Then $A \subset W$ and $W - A \subset V_i - A$. Thus

$$\mathcal{H}^{n-m}\big((W - A) \cap f^{-1}(y)\big) \le \mathcal{H}^{n-m}\big((V_i - A) \cap f^{-1}(y)\big) \to 0 \text{ as } i \to \infty$$

Hence, for $\mathcal{L}^m$ a.e $y \in \mathbb{R}^m$,

$$\mathcal{H}^{n-m}\Big((W - A) \cap f^{-1}(y)\Big) = 0$$

Therefore, for $\mathcal{L}^m$ a.e $y \in \mathbb{R}^m$,$y \to \big(W - A\big) \cap f^{-1}(y)$ is $\mathcal{H}^{n-m}$ measurable. Now,

$$A \cap f^{-1}(y) = \big(W \cap f^{-1}(y)\big) - \Big((W - A) \cap f^{-1}(y)\Big) \text{ and } W \cap f^{-1}(y) \text{ is a borel set.}$$

Hence, for $\mathcal{L}^m$ a.e $y \in \mathbb{R}^m$, $A \cap f^{-1}(y)$ is $\mathcal{H}^{n-m}$ measurable. This proves the claim 3.

Case 4 :

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

Write

$$A = \bigcup_{k=1}^{\infty} A_k$$

$$\text{where } A_k = A \cap B(0, k) \quad , \quad k \in \mathbb{N}$$

Apply case 3 to A_k to get:
$A \cap f^{-1}(y) = \cup_{i=1}^{\infty} A_k \cap f^{-1}(y)$ and hence, it is $\mathcal{H}^{n-m}$ measurable for $\mathcal{L}^m$ almost every y. Since $A_k \subset A_{k+1}$ and $\mathcal{H}^{n-m}$ is borel regular,

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

And thus, $y \to \mathcal{H}^{n-m}\big(A \cap f^{-1}(y)\big)$ is $\mathcal{L}^m$ measurable.

$\square$

Remark 2.72. The same calculations also show that for each $A \subset \mathbb{R}^n$,

$$\int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap f^{-1}(y)\big) d\mathcal{H}^l \leq \frac{\alpha(k)\alpha(l)}{\alpha(k+l)} \big(Lip(f)\big)^l \mathcal{H}^{k+l}(A)$$

Lemma 2.73. *Let*

$$t > 1$$

Assume that $h : \mathbb{R}^n \to \mathbb{R}^n$, and set

$$B := \{x \ ; \ Dh(x) \ exists \ \ , \ Jh(x) > 0\}$$

Then, there are countable collection of disjoint borel subsets of $\mathbb{R}^n$,say $\{D_k\}_{k=1}^{\infty}$, such that

- $\mathcal{L}^n(B - \bigcup_{k=1}^{\infty} D_k) = 0$
- *h is injective on D_k for $k \in \mathbb{N}$.*
- *For each $k \in \mathbb{N}$, there exists symmetric automorphism $S_k : \mathbb{R}^n \to \mathbb{R}^n$, such that*
  - $Lip\big(S_k^{-1} \circ (h|_{D_k})\big) \leq t$
  - $Lip\big((h|_{D_k})^{-1} \circ S_k\big) \leq t$
  - $\frac{1}{t^n}|detS_k| \leq Jh|_{D_k} \leq t^n|detS_k|$

Proof. By the lemma 2.67, with h in place of f, we have existence of disjoint borel sets $\{E_k\}_{k\in\mathbb{N}}$ and symmetric automorphisms $\{T_k : \mathbb{R}^n \to \mathbb{R}^n\}_{k\in\mathbb{N}}$ such that

- $B = \bigcup_{k=1}^{\infty} E_k$
- h is injective on E_k.
- $Lip(h|_{E_k} \circ T_k^{-1}) \leq t$
- $Lip(T_k \circ h|_{E_k}^{-1}) \leq t$
- $\frac{1}{t^n}|detT_k| \leq Jh|_{E_k} \leq t^n|detT_k|$

Observe that the above also tells that $h|_{E_k}^{-1}$ is lipshitz and thus by the lipshitz extension theorem, there is lipshitz mapping $h_k : \mathbb{R}^n \to \mathbb{R}^n$ such that, on $h(E_k)$, $h_k \equiv \big(h|_{E_k}^{-1}\big)$.
Claim 1 : $Jh_k > 0$ for $\mathcal{L}^n$ almost every point in $h(E_k)$.
Proof of the claim 1 : Observe that on E_k ,

$$h_k \circ h = \text{identity}$$

By the corollary of the Rademacher's theorem, we have

$$Dh_k\big(h(x)\big) \circ Dh(x) = I \text{ for } \mathcal{L}^n \text{ almost every } x \text{ on } E_k$$

$$\implies Jh_k\big(h(x)\big)Jh(x) = 1 \text{ for } \mathcal{L}^n \text{ almost every } x \text{ on } E_k$$

This tells that $Jh_k\big(h(x)\big) > 0$ for $\mathcal{L}^n$ a.e $x \in E_k$. Since the function h is lipshitz, the claim follows.
Applying the lemma 2.67 for h_k, we have the existence of disjoint borel sets $\{F_{jk}\}_{j=1}^{\infty}$ and symmetric automorphisms $\{R_{jk}\}_{j=1}^{\infty}$ such that

- $\mathcal{L}^n\Bigg(h(E_k) - \bigcup_{j=1}^{\infty} F_{jk}\Bigg) = 0$
- h_k is injective on F_{jk}
- $Lip\big((h_k|_{F_{jk}}) \circ R_{jk}^{-1}\big) \leq t$
- $Lip\big(R_{jk} \circ h_k|_{F_{jk}}^{-1}\big) \leq t$
- $\frac{1}{t^n}|detR_{jk}| \leq Jh_k|_{F_{jk}} \leq t^n|detR_{jk}|$

Declare

$$D_{jk} := E_k \cap h^{-1}\big(F_{jk}\big)$$

$$S_{jk} := R_{jk}^{-1}$$

Claim 2 :

$$\mathcal{L}^n\Bigg(B - \bigcup_{j,k=1}^{\infty} D_{jk}\Bigg) = 0$$

Proof of the claim 2 : Observe that

$$h_k\Bigg(h(E_k) - \bigcup_{j=1}^{\infty} F_{jk}\Bigg) = h^{-1}\Bigg(h(E_k) - \bigcup_{j=1}^{\infty} F_{jk}\Bigg) = E_k - \bigcup_{j=1}^{\infty} D_{jk}$$

Thus the by the construction, we have

$$\mathcal{L}^n\Bigg(E_k - \bigcup_{j=1}^{\infty} D_{jk}\Bigg) = 0$$

This proves the claim 2.
By the construction, clearly, h is injective on D_{jk}.
Claim 3 : For $j, k \in \mathbb{N}$,

- $Lip\big(S_{jk}^{-1} \circ h|_{Djk}\big) \leq t$
- $Lip\big(h|_{D_{jk}}^{-1} \circ S_{jk}\big) \leq t$
- $\frac{1}{t^n}|detS_{jk}| \leq Jh|_{D_{jk}} \leq t^n|detS_{jk}|$

Proof of the claim 3 :

$$Lip\big(S_{jk}^{-1} \circ h|_{D_{jk}}\big) = Lip\big(R_{jk} \circ h|_{D_{jk}}\big) \leq Lip\big(R_{jk} \circ h_k|_{F_{jk}}^{-1}\big) \leq t$$

Similarly the other lipshitz inequality follows. Also, as noted above,

$$Jh_k\big(h(x)\big)Jh(x) = 1 \ \mathcal{L}^n \text{ almost everywhere on } D_{jk}$$

$$\implies \frac{1}{t^n}|detS_{jk}| = \frac{1}{t^n}|detR_{jk}|^{-1} \leq Jh|_{Djk} \leq t^n|detR_{jk}|^{-1} = t^n|detS_{jk}|$$

This proves the claim 3 and hence the lemma. □

Theorem 2.74. ***Co-Area formula :***
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be lipshitz mapping, $n \geq m$. Then for each $\mathcal{L}^n$ measurable set $A \subset \mathbb{R}^n$,

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

Proof. Again, as per the Rademacher's theorem, we can assume that $Df(x)$ and thus $Jf(x)$ exists for all $x \in A$ and also that $\mathcal{L}^n(A) < \infty$. As before, splitting the proof into 2 cases :
Case 1 :

$$A \subset \{Jf > 0\}$$

For each $\lambda \in \Lambda(n, n-m)$ write $f = q \circ h_\lambda$, where

$$h_\lambda : \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^{n-m} \ ; \ x \to \big(f(x), P_\lambda(x)\big)$$

$$q : \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m \ ; \ (y, z) \to y$$

Declare

$$A_\lambda := \{x \in A \ ; \ det\big(Dh_\lambda\big) \neq 0\} \equiv \{x \in A \ ; \ P_\lambda|_{\big(Df(x)\big)^{-1}(0)} \text{ is one-one}\}$$

Now,

$$A = \bigcup_{\lambda \in \Lambda} A_\lambda$$

Therefore, assume for now that A is some A_λ. Fix $t > 1$ and apply the previous lemma to h_λ to get disjoint borel sets $\{D_k\}$ and symmetric automorphisms $\{S_k\}$ as in the hypothesis of the previous lemma. Set

$$G_k := A \cap D_k$$

Claim 1 :

$$\frac{1}{t^n}[[q \circ S_k]] \leq Jf|_{G_k} \leq t^n[[q \circ S_k]]$$

Proof of the claim 1 : Since $f = q \circ h$, for $\mathcal{L}^n$ almost everywhere,

$$Df = q \circ Dh = q \circ S_k \circ S_k^{-1} \circ Dh = q \circ S_k \circ D\big(S_k^{-1} \circ h\big) = q \circ S_k \circ C$$

where

$$C := D\big(S_k^{-1} \circ h\big)$$

By the previous lemma, on G_k,

$$\frac{1}{t} \leq Lip\big(S_k^{-1} \circ h\big) = Lip(C) \leq t$$

Declare the following by the polar decomposition:

$$Df = S \circ O^*$$

$$q \circ S_k = T \circ P^*$$

where

$$S, T : \mathbb{R}^m \to \mathbb{R}^m \text{ symmetric}$$

$$O, P : \mathbb{R}^m \to \mathbb{R}^n \text{ orthogonal}$$

Then

$$S \circ O^* = T \circ P^* \circ C$$

As a consequence,

$$S = T \circ P^* \circ C \circ O$$

By the assumption of the case $G_k \subset \{Jf > 0\}$, $detS \neq 0 \implies detT \neq 0$. Hence, for $v \in \mathbb{R}^m$,

$$|T^{-1} \circ Sv| = |P^* \circ C \circ O(v)| \leq |C \circ O(v)| \leq t|O(v)| = t|v|$$

Thus

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

$$\implies Jf = |detS| \leq t^n |detT| = t^n [[q \circ S_k]]$$

Similarly, we have the other inequality :

$$[[q \circ S_k]] = |detT| \leq t^n |detS| = t^n Jf$$

Now,

$$t^{-3n+m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\bigg(G_k \cap f^{-1}(y)\bigg) dy$$

$$= t^{-3n+m}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(h^{-1}\big(h(G_k)\big)\cap q^{-1}(y)\Big)dy$$

$$\le t^{-2n}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(S_k^{-1}\big(h(G_k)\cap q^{-1}(y)\big)\Big)dy$$

$$= t^{-2n}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(S_k^{-1}\circ h(G_k)\cap \big(q\circ S_k\big)^{-1}(y)\Big)dy$$

By the lemma 2.70,

$$t^{-2n}[[q\circ S_k]]\mathcal{L}^n\Big(S_k^{-1}\circ h(G_k)\Big)$$

$$\le t^{-n}[[q\circ S_k]]\mathcal{L}^n(G_k)$$

$$\le \int_{G_k} Jf dx$$

Similar proof shows that

$$\int_{G_k} Jf dx \le t^{3n-m}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(G_k\cap f^{-1}(y)\Big)dy$$

Using the fact that A is same as the set $\bigcup G_k$ in $\mathcal{L}^n$ measure, summing on k and letting $t\to 1+$, we have

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

Case 2 :

$$A\subset\{Jf=0\}$$

Fix $0<\epsilon\le 1$. Define

$$g:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m\ ,\ (x,y)\to f(x)+\epsilon y$$

$$p:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m\ ,\ (x,y)\to y$$

Then,

$$Dg = \begin{bmatrix} Df & \epsilon I\end{bmatrix}_{m\times(n+m)}$$

Also, $\epsilon^m\le Jg = [[Dg]] = [[Dg^*]]\le C\epsilon$.

Note that

$$\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(A\cap f^{-1}(y)\Big)dy$$

$$= \int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(A\cap f^{-1}\{y-\epsilon w\}\Big)dy\quad,\quad \forall w\in\mathbb{R}^n$$

$$= \frac{1}{\alpha(m)}\int_{B(0,1)}\int_{\mathbb{R}^m}\mathcal{H}^{n-m}\Big(A\cap f^{-1}\{y-\epsilon w\}\Big)dydw$$

Claim 2 : Fix $y \in \mathbb{R}^m$, $w \in \mathbb{R}^m$. Set

$$B := A \times B(0,1) \subset \mathbb{R}^{n+m}$$

Then

$$B \cap g^{-1}(y) \cap p^{-1}(w) = \begin{Bmatrix} \phi & w \notin B(0,1) \\ \big(A \cap f^{-1}\{y - \epsilon w\big) \times \{w\} & w \in B(0,1) \end{Bmatrix}$$

Proof of the claim 2 :

$$(x,z) \in B \cap g^{-1}(y) \cap p^{-1}(w) \iff x \in a, z \in B(0,1), f(x) + \epsilon z = y, z = w$$

$$\iff x \in A, z = w \in B(0,1), f(x) = y - \epsilon w$$

$$\iff w \in B(0,1), (x,z) \in \big(A \cap f^{-1}\{y - \epsilon w\}\big) \times \{w\}$$

This proves the claim 2 .

Now, the intergral equality becomes,

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

$$= \frac{1}{\alpha(m)} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\big(B \cap g^{-1}\{y\} \cap p^{-1}\{w\}\big) dw \ dy$$

By the remark 2.72,

$$\leq \frac{\alpha(n-m)}{\alpha(n)} \int_{\mathbb{R}^m} \mathcal{H}^n\big(B \cap g^{-1}\{y\}\big) dy$$

By the case 1,

$$= \frac{\alpha(n-m)}{\alpha(m)} \int_B Jg dx \ dz$$

$$\leq \frac{\alpha(n-m)\alpha(m)}{\alpha(n)} \mathcal{L}^n(A) \sup_B Jg$$

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

Now $\epsilon > 0$ was arbitrary and hence the theorem holds for the case 2. For the general case, split the set A as $A_1 \cup A_2$, where $A_1 \subset \{Jf > 0\}$ and $A_2 \subset \{Jf = 0\}$ and apply the above 2 cases to conclude the theroem. □

Theorem 2.75. *Change of Variables :*
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be lipshitz and $n \geq m$. Then , for all $g : \mathbb{R}^n \to \mathbb{R}$, integrable function,

$$g|_{f^{-1}\{y\}} \textit{ is } \mathcal{H}^{n-m} \textit{ integrable for } \mathcal{L}^m \textit{ almost every } y \in \mathbb{R}^m$$

$$\int_{\mathbb{R}^n} g(x) Jf(x) dx = \int_{\mathbb{R}^m} \left(\int_{f^{-1}\{y\}} g d\mathcal{H}^{n-m} \right) dy$$

Proof. Let $g := \sum_{i=1}^{n} C_i \mathcal{X}_{A_i}$ be a simple function. Then

$$\int_{\mathbb{R}^n} g\ Jf\ dx = \sum_{i=1}^{n} C_i \int_{A_i} Jf\ dx$$

$$= \sum_{i=1}^{n} C_i \int_{\mathbb{R}^m} \mathcal{H}^{n-m}\Big(A_i \cap f^{-1}\{y\}\Big)\ dy$$

$$= \int_{\mathbb{R}^m} \sum_{i=1}^{n} C_i \mathcal{H}^{n-m}\Big(A_i \cap f^{-1}\{y\}\Big)\ dy$$

$$\int_{\mathbb{R}^m} \Big(\int_{f^{-1}\{y\}} g d\mathcal{H}^{n-m}\Big) dy$$

Since this is true for simple functions, as done in the area's change of variables, by DCT and MCT, one can conclude the theorem for all integrable functions. □

2.5.2 Applications

- Polar Co-ordinates :

  Let $g : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{L}^n$ integrable function. Then

  $$\int_{\mathbb{R}^n} g\ dx = \int_0^\infty \Big(\int_{\partial B(0,r)} g d\mathcal{H}^{n-1}\Big) dr$$

  In particular,

  $$\frac{d}{dr}\Big(\int_{B(0,r)} g\ dx\Big) = \int_{\partial B(0,r)} g d\mathcal{H}^{n-1}$$

  Proof. Declare

  $$f(x) := |x| \equiv \sqrt{\sum_{i=1}^{n} x_i^2}$$

  Then,

  $$Df(x) = \frac{x}{|x|}\ ;\ Jf(x) = 1$$

  The change of variables formula concludes the above statement. □

- Level Sets :

  Let $f : \mathbb{R}^n \to \mathbb{R}$ be lipshitz. Then,

  $$\int_{\mathbb{R}^n} |Df| dx = \int_{\mathbb{R}} \mathcal{H}^{n-1}\big(\{f = t\}\big) dt$$

  This follows form the fact that $Jf = |Df|$.

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be lipshitz with $ess\ \inf |Df| > 0$. Assume that $g : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{L}^n$ integrable. Then,

$$\int_{\{f>t\}} g dx = \int_t^\infty \left(\int_{\{f=s\}} \frac{g}{|Df|} d\mathcal{H}^{n-1} \right) ds$$

In particular,

$$\frac{d}{ds} \int_{\{f>s\}} f dx = - \int_{\{f=s\}} \frac{g}{|Df|} d\mathcal{H}^{n-1} \text{ for } \mathcal{L}^1 \text{ almost every } s$$

Proof. As before, let

$$Jf = |Df|$$

Declare

$$E_s := \{f > s\}$$

By the Co-Area formula,

$$\int_{\{f>s\}} g dx = \int_{\mathbb{R}^n} \mathcal{X}_{E_s} \frac{g}{|Df|} Jf dx$$

$$= \int_{\mathbb{R}} \left(\int_{\partial E_t} \frac{g}{|Df|} \mathcal{X}_{E_t} d\mathcal{H}^{n-1} \right) dt$$

$$= \int_t^\infty \left(\int_{\partial E_t} \frac{g}{|Df|} \mathcal{X}_{E_t} d\mathcal{H}^{n-1} \right) dt$$

□

References

- Measure Theory and Fine Properties of Functions by Lawrence C. Evans and Ronald F.Gariepy, CRC Press.

- Real and Complex Analysis by Walter Rudin, McGraw - Hill International Editions.

- Functional Analysis by Walter Rudin, Tata-Mcgraw Hill Publications.

- Real Analysis by Elias M.Stein and Rami Shakarchi, Princeton University Press.

- Real Analysis by H.L. Royden, Macmillan Publishing Company, Collier Macmillan Publishing.

- Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory by Francesco Maggi, Cambridge Studies in Advanced Mathematics.

- TIFR Lecture notes on Analysis by Raghavan Narasimhan.