Measure Theory and Lebesgue Integration

T. Muthukumar
tmk@iitk.ac.in

April 12, 2018

Contents

Notations

Symbols

2^S will denote the power set, the set of all subsets, of a set S

$\mathcal{L}(\mathbb{R}^n)$ the class of all Lebesgue measurable subsets of $\mathbb{R}^n$

$\mathbb{C}$ denotes the plane of complex numbers

$\mathbb{Q}$ denotes the set of all rationals in $\mathbb{R}$

$\mathbb{Q}^n$ set of all vectors in $\mathbb{R}^n$ with each coordinate being rational number

$\mathbb{R}$ denotes the real line

$\mathbb{R}^n$ denotes the Euclidean space of dimension n

Function Spaces

$\mathcal{R}([a,b])$ denotes the space of all Riemann integrable functions on the interval $[a,b]$

$\mathrm{Lip}(E)$ denotes the space of all Lipschitz functions on E

$AC(E)$ denotes the space of all absolutely continuous functions on E

$BV(E)$ denotes the space of all bounded variation functions on E

$C(X)$ the class of all real-valued continuous functions on X

$C_0(X)$ denotes the space of all continuous functions vanishing at ∞ on X

$C_c(X)$ denotes the space of all compactly supported continuous functions on X

$L^p(E)$ denotes the space of all measurable p-integrable functions on E

$M(\mathbb{R}^n)$ the class of all finite a.e. real valued Lebesgue measurable functions on $\mathbb{R}^n$

General Conventions

$\overline{B}_r(x)$ will denote the closed ball of radius r and centre at x

E^c will denote the set complement of $E \subset S$, $S \setminus E$

Chapter 1

Introduction

1.1 Riemann Integration and its Inadequacy

Let $f : [a, b] \to \mathbb{R}$ be a *bounded* function. Let P be the partition of the interval $[a, b]$, $a = x_0 \leq x_1 \leq \ldots \leq x_k = b$. For $i = 0, 1, 2, \ldots, k$, let

$$M_i(P) = \sup_{x \in [x_{i-1}, x_i]} f(x) \text{ and } m_i(P) = \inf_{x \in [x_{i-1}, x_i]} f(x).$$

The upper Riemann sum of f with respect to the given partition P is,

$$U(P, f) = \sum_{i=1}^{k} M_i(P)(x_i - x_{i-1})$$

and the lower Riemann sum of f with respect to the given partition P is,

$$L(P, f) = \sum_{i=1}^{k} m_i(P)(x_i - x_{i-1}).$$

We say the bounded function f is *Riemann integrable* on $[a, b]$ if the infimum of upper sum and supremum of lower sum, over all partitions P of $[a, b]$, coincide and is denoted as

$$\int_a^b f(x)\, dx := \inf_P U(P, f) = \sup_P L(P, f).$$

If $f = u + iv$ is a bounded complex-valued function on $[a, b]$, then f is said to be Riemann integrable if its real and imaginary parts are Riemann integral and

$$\int_a^b f(x)\,dx = \int_a^b u(x)\,dx + i\int_a^b v(x)\,dx.$$

If either f is *unbounded* or the domain $[a,b]$ is *not* finite then its corresponding integral, called as *improper integral*, is defined in terms of limits of Riemann integrable functions, whenever possible.

Exercise 1. Every Riemann integrable function[1] is bounded.

Let $\mathcal{R}([a,b])$ denote the space of all Riemann integrable functions on $[a,b]$. The space $\mathcal{R}([a,b])$ forms a vector space over $\mathbb{R}$ (or $\mathbb{C}$). It is closed under composition, if it makes sense.

Theorem 1.1.1. *If f is continuous on $[a,b]$, then $f \in \mathcal{R}([a,b])$.*

In fact even piecewise continuity is sufficient for Riemann integrability.

Theorem 1.1.2.

If f is continuous except at finitely many points of $[a,b]$ (piecewise continuous), then $f \in \mathcal{R}([a,b])$.

But there are functions which has discontinuity at countably many points and are still in $\mathcal{R}([a,b])$.

Example 1.1. Consider the function

$$f(x) = \begin{cases} 1 & \text{if } \frac{1}{k+1} < x \le \frac{1}{k} \text{ and } k \text{ is odd} \\ 0 & \text{if } \frac{1}{k+1} < x \le \frac{1}{k} \text{ and } k \text{ is even} \\ 0 & x = 0 \end{cases}$$

which has discontinuities at $x = 0$ and $x = 1/k$, for $k = 1, 2, \ldots$. It can be shown that $f \in \mathcal{R}([0,1])$.

Theorem 1.1.3.

If f is bounded monotonic on $[a,b]$ then $f \in \mathcal{R}([a,b])$.

In fact, one can construct functions whose set of discontinuities are 'dense' in $[0,1]$.

[1] here by Riemann integrable we mean the upper sum and lower sum coincide and are finite

Example 1.2. For instance, let $\{r_k\}_1^\infty$ denote a countable dense subset of $[0,1]$ (for instance, $\mathbb{Q}$) and define

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k^2} H(x - r_k)$$

where $H : \mathbb{R} \to \mathbb{R}$ is defined as

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

The function f is discontinuous at all the points r_k and can be shown to be in $\mathcal{R}([0,1])$, because it is bounded and monotone.

Theorem 1.1.4. *If $f \in \mathcal{R}([a,b])$ then f is continuous on a dense subset of $[a,b]$.*

Example 1.3. An example of a function $f : [0,1] \to \mathbb{R}$ which is not Riemann integrable is

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

A necessary and sufficient condition of Riemann integrability is given by Theorem 3.0.1. Thus, even to characterise the class of Riemann integrable functions, we need to have the notion of length ("measure") (at least measure zero).

1.1.1 Limit and Integral: Interchange

Let us consider a sequence of functions $\{f_k\} \subset \mathcal{R}([a,b])$ and define $f(x) := \lim_{k\to\infty} f_k(x)$, assuming that the limit exists for every $x \in [a,b]$. Does $f \in \mathcal{R}([a,b])$? The answer is a "no", as seen in example below.

Example 1.4. Fix an enumeration (order) of the set of rationals in $[0,1]$. Let the finite set r_k denote the first k elements of the set of rationals in $[0,1]$. Define the sequence of functions

$$f_k(x) = \begin{cases} 1 & \text{if } x \in r_k \\ 0 & \text{otherwise.} \end{cases}$$

Each $f_k \in \mathcal{R}([0,1])$, since it has discontinuity at k (finite) number of points. The point-wise limit of f_k, $f = \lim_{k\to\infty} f_k$, is

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

which we have seen above is not Riemann integrable.

Thus, the space $\mathcal{R}([a,b])$ is not "complete" under point-wise limit. However, $\mathcal{R}([a,b])$ is complete under uniform convergence.

A related question is if the limit $f \in \mathcal{R}([a,b])$, is the Riemann integral of f the limit of the Riemann integrals of f_k, i.e., can we say

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} \int_a^b f_k(x)\,dx?$$

The answer is a "no" again.

Example 1.5. Consider the functions

$$f_k(x) = \begin{cases} k & x \in (0,1/k) \\ 0 & \text{otherwise} \end{cases}$$

Then $f(x) = \lim_k f_k(x) = 0$. Note that

$$\int_{\mathbb{R}} f_k(x)\,dx = 1 \quad \forall k,$$

but $\int_{\mathbb{R}} f(x)\,dx = 0$.

The interchange becomes possible under uniform convergence.

Theorem 1.1.5. *Let $\{f_k\} \subset \mathcal{R}([a,b])$ and $f_k(x) \to f(x)$ uniformly in $[a,b]$. Then $f \in \mathcal{R}([a,b])$ and*

$$\int_a^b f = \lim_{k\to\infty} \int_a^b f_k.$$

But uniform convergence is too demanding in practice. The following more general result for interchanging limit and integral will be proved in this write-up.

Theorem 1.1.6. *Let* $\{f_k\} \subset \mathcal{R}([a,b])$ *and* $f \in \mathcal{R}([a,b])$*. Also, let* $f_k(x) \to f(x)$ *point-wise and* f_k *are uniformly bounded. Then*

$$\lim_{k\to\infty} \int_a^b f_k = \int_a^b f.$$

The proof of above theorem is not elementary, thus in classical analysis we always prove the result for uniform convergence. Observe the hypothesis of integrability on f in the above theorem.

1.1.2 Differentiation and Integration: Duality

An observation we make, once we have Riemann integration, is about the dual nature of differentiation and integration. Thus, one asks the following two questions:

1. (Derivative of an integral) For which class of functions can we say
$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)?$$
2. (Integral of a derivative) For which class of functions can we say
$$\int_a^b f'(x)\,dx = f(b) - f(a)?$$

To answer the first question, for any $f \in \mathcal{R}([a,b])$, let us define the function

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

Exercise 2. Show that if $f \in \mathcal{R}([a,b])$ then F is continuous on $[a,b]$.

The first question is answered by the following result of Riemann integration.

Theorem 1.1.7. *Let* $f \in \mathcal{R}([a,b])$*. If* f *is continuous at a point* $x \in [a,b]$*, then* F *is differentiable at* x *and* $F'(x) = f(x)$*.*

What is the most general class of functions for which the above result holds true.

The second question is answered by the famous Fundamental theorem of calculus (FTC).

Theorem 1.1.8 (Fundamental Theorem of Calculus). *If f is differentiable function on $[a, b]$ such that $f' \in \mathcal{R}([a, b])$, then*

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

Note that the fundamental theorem of calculus fails under the following two circumstances:

1. For a continuous function f on $[a, b]$ which is nowhere differentiable on $[a, b]$. Do such functions exist?

2. Derivative of f exists for all points in $[a, b]$, but f' is not integrable. Do such functions exist?

K. Weierstrass was the first to show in 1872 the existence of a everywhere continuous function which is nowhere differentiable. Prior to Weierstrass' proof it was believed that every continuous function is differentiable except on a set of "isolated" points. This example of Weierstrass showed the existence of function for which FTC may not make any sense.

The existence of a function f whose derivative exists everywhere but the derivative is not integrable, was shown by Vito Volterra, who was a student of Ulisse Dini, in 1881. His example was a clever modification of the function

$$g(x) = \begin{cases} x^2 \sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

which is differentiable. The derivative of g, $g'(x) = 2x \sin(1/x) - \cos(1/x)$, is discontinuous at $x = 0$.

A natural question to ask was: Identify the class of functions for which FTC makes sense.

1.2 Motivating Lebesgue Integral and Measure

Whatever the reasons are, we should be convinced now that it is worthwhile looking for a new type of integration which coincides for Riemann integrable functions and also includes "non-integrable" (Riemann) functions.

The Riemann integration was based on the simple fact that one can integrate step functions (piecewise constant) and then approximate any given function with piecewise constant functions, by partitioning the *domain* of the function. Lebesgue came up with this idea of partitioning the *range* of the function.

A very good analogy to motivate Lebesgue integration is the following (cf. [Pug04]): Suppose A asks both B and C to give the total value for a bunch of coins with all denominations lying on a table. First B counts them as he picks the coins and adds their denomination to come up with the total value. This is Riemann's way of integration (partitioning the domain, if you consider the function to be coin mapped to its denomination). In his/her turn, C sorts the coin as per their denominations in to separate piles and counts the coins in each pile, multiply it with the denomination of the pile and sum them up for the total value. Both B and C will come up with the same value (assuming they counted right!). The way C counted is Lebesgue's way of integration.

We know that integration is related with the question of computing length/area/volume of a subset of an Euclidean space, depending on its dimension. Now, if one wants to partition the range of a function, we need some way of "measuring" how much of the domain is sent to a particular region of the partition. This problem leads us to the theory of measures where we try to give a notion of "measure" to subsets of an Euclidean space.

Chapter 2

Lebesgue Measure on $\mathbb{R}^n$

2.1 Introduction

In this chapter we shall develop the notion of Lebegue 'measure' in $\mathbb{R}^n$.

Definition 2.1.1. *We say R is a* cell *(open) in $\mathbb{R}^n$ if R is of the form $(a_1, b_1) \times (a_2, b_2) \times \ldots \times (a_n, b_n)$, i.e.,*

$$R = \Pi_{i=1}^n (a_i, b_i) := \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \in (a_i, b_i) \text{ for all } 1 \leq i \leq n\}.$$

The volume *(finite) of the cell R, denoted as $|R|$, is the non-negative number, $|R| = \Pi_{i=1}^n (b_i - a_i)$. We say R is closed if $R = \Pi_{i=1}^n [a_i, b_i]$.*

We do not let $a_i = -\infty$ or $b_i = +\infty$, i.e., by a cell we always refer to a cell with finite volume. In fact, by our definition, cells in $\mathbb{R}^n$ are precisely those rectangles with finite volume in $\mathbb{R}^n$ whose sides are parallel to the coordinate axes. Also, a cell could refer to Cartesian products of open, closed or half-open or half-closed intervals.

Note that if $a_i = b_i$, for all i, then we have the volume of an empty set to be zero. Moreover, if $\overline{R}$ is the closure of an open cell R, then $|\overline{R}| = |R|$. If $a_i = b_i$, for some i in a closed cell, then the (lower dimensional) cell also has volume zero.

Exercise 3. Show that the volume of a cell is translation invariant, i.e., for any cell R, $|R + x| = |R|$ for all $x \in \mathbb{R}^n$. The volume is also dilation invariant, i.e., for any $\lambda = (\lambda_1, \ldots, \lambda_n)$ such that $\lambda_i > 0$, $|\lambda R| = \prod_{i=1}^n \lambda_i |R|$, where $\lambda R = \{(\lambda_i x_i) \mid (x_i) \in R\}$.

Exercise 4 (uniqueness of volume). Let $\mathcal{C}$ be the collection of all cells of $\mathbb{R}^n$ and if $\nu : \mathcal{C} \to [0, +\infty)$ is a well defined set-function on $\mathcal{C}$ such that ν is invariant under translation and dilation. Show that ν is same as the volume $|\cdot|$ up to a constant, i.e., there exists a constant $\alpha \geq 0$ such that $\nu(R) = \alpha|R|$ for all cells $R \in \mathcal{C}$. In particular, if we additionally impose the condition that $\nu([0,1]^n) = 1$ then $\nu(R) = |R|$ for all $R \in \mathcal{C}$.

Exercise 5. Show that the volume satisfies *monotonicity*, i.e., if $R \subset Q$, then $|R| \leq |Q|$.

Exercise 6. If $\{R_i\}_1^k$ are cells in $\mathbb{R}^n$ such that $R \subset \cup_{i=1}^k R_i$, then $|R| \leq \sum_{i=1}^k |R_i|$.

Theorem 2.1.2. *For every open subset $\Omega \subset \mathbb{R}$, there exists a unique countable family of open intervals I_i such that $\Omega = \cup_{i=1}^\infty I_i$ where I_i's are pairwise disjoint.*

Proof. Since Ω is open, for every $x \in \Omega$, there is an open interval in Ω that contains x. Let us pick the largest such open interval in Ω that contains x. How do we do this? Let, for each $x \in \Omega$,

$$a_x := \inf_{a<x}\{(a,x) \subset \Omega\} \quad \text{and } b_x := \sup_{b>x}\{(x,b) \subset \Omega\}.$$

Of course, a_x and b_x can take $\pm\infty$. Note that $a_x < x < b_x$. Set $I_x := (a_x, b_x)$, is the largest open interval in Ω containing x. Thus, we have $\Omega = \cup_{x\in\Omega} I_x$. We shall now note that for any $x, y \in \Omega$ such that $x \neq y$, either $I_x = I_y$ or $I_x \cap I_y = \emptyset$. Suppose $I_x \cap I_y \neq \emptyset$ then $I_x \cup I_y$ is also an open interval in Ω that contains x. Therefore, by the maximality of I_x, $I_x \cup I_y \subset I_x$. Hence, $I_x = I_x \cup I_y$. Similarly, $I_x \cup I_y = I_y$. Thus, $I_x = I_y$ and Ω is a disjoint union of open intervals. It now only remains to show that the union can be made countable. Note that every open interval I_x contains a rational number. Since different intervals are disjoint, we can pick distinct rationals from each interval. Since rationals are countable, the collection of disjoint intervals cannot be uncountable. Thus, we have a countable collection of disjoint open intervals I_i such that $\Omega = \cup_{i=1}^\infty I_i$. □

From the uniqueness in the result proved above we are motivated to define the "length" of an open set $\Omega \subset \mathbb{R}$ as the sum of the lengths of the intervals I_i. But this result has no exact analogue in $\mathbb{R}^n$, for $n \geq 2$.

Exercise 7. An open connected set $\Omega \subset \mathbb{R}^n$, $n \geq 2$ is the disjoint union of open cells iff Ω is itself an open cell.

Exercise 8. Show that an open disc in $\mathbb{R}^2$ cannot be the disjoint union of open cells.

However, relaxing our requirement to almost disjoint-ness (defined below) will generalise Theorem 2.1.2 to higher dimensions $\mathbb{R}^n$, $n \geq 2$.

Definition 2.1.3. *We say a collection of cells R_i to be* almost disjoint *if the interiors of R_i are pairwise disjoint.*

Exercise 9. If a cell $R = \cup_{i=1}^k R_i$ such that R_i are pairwise almost disjoint, then $|R| = \sum_{i=1}^k |R_i|$.

Theorem 2.1.4. *For every open subset $\Omega \subset \mathbb{R}^n$, there exists a countable family of almost disjoint closed cells R_i such that $\Omega = \cup_{i=1}^\infty R_i$.*

Proof. To begin we consider the grid of cells in $\mathbb{R}^n$ of side length 1 and whose vertices have integer coordinates. The number cells in the grid is countable and they are almost disjoint. We ignore all those cells which are contained in Ω^c. Now we have two families of cells, those which are contained in Ω, call the collection $\mathcal{C}$, and those which intersect both Ω and Ω^c. We bisect the latter cells further in to 2^n cells of side each $1/2$. Again ignore those contained in Ω^c and add those contained in Ω to the collection $\mathcal{C}$. Further bisecting the common cells in to cells of side length $1/4$. Repeating this procedure, we have a countable collection $\mathcal{C}$ of almost disjoint cells in Ω. By construction, $\cup_{R \in \mathcal{C}} R \subset \Omega$. Let $x \in \Omega$ then there is a cell of side length $1/2^k$ (after bisecting k times) in $\mathcal{C}$ which contains x. Thus, $\cup_{R \in \mathcal{C}} R = \Omega$. $\square$

Again, as we did in one dimension, we hope to define the "volume" of an open subset $\Omega \subset \mathbb{R}^n$ as the sum of the volumes of the cells R obtained in above theorem. However, since the collection of cells is not unique, in contrast to one dimension, it is not clear if the sum of the volumes is independent of the choice of your family of cells.

We wish to extend the notion of volume to arbitrary subsets of an Euclidean space such that they coincide with the usual notion of volume for a cell, most importantly, preserving the properties of the volume. So, what are these properties of volume we wish to preserve? To state them, let's first regard the volume as a set function on the power set of $\mathbb{R}^n$, mapping to a non-negative real number. Thus, we wish to construct a 'measure' μ, $\mu : 2^{\mathbb{R}^n} \to [0, \infty]$ such that

1. If R is any cell of $\mathbb{R}^n$, then $\mu(R) = |R|$.

2. (**Translation Invariance**) For every $E \subset \mathbb{R}^n$, $\mu(E + x) = \mu(E)$ for all $x \in \mathbb{R}^n$.

3. (**Monotonicity**) If $E \subset F$, then $\mu(E) \le \mu(F)$.

4. (**Countable Sub-additivity**) If $E = \cup_{i=1}^{\infty} E_i$ then $\mu(E) \le \sum_{i=1}^{\infty} \mu(E_i)$.

5. (**Countable Additivity**) If $E = \cup_{i=1}^{\infty} E_i$ such that E_i are pairwise disjoint then $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$.

Exercise 10. Show that if μ obeys finite additivity and is non-negative, then μ is monotone. (Basically monotonicity is redundant from countable additivity).

2.2 Outer measure

To construct a 'measure' on the power set of $\mathbb{R}^n$, we use the simple approach of 'covering' an arbitrary subset of $\mathbb{R}^n$ by cells and assigning a unique number using them.

Definition 2.2.1. *Let $E \subseteq \mathbb{R}^n$, a subset of $\mathbb{R}^n$. We say that a family of cells $\{R_i\}_{i \in I}$ is a* cover *of E iff $E \subseteq \cup_{i \in I} R_i$. If each of the cell R_i in the cover is an open (resp. closed) cell, then the cover is said to be* open (resp. closed) cover *of E. If the index set I is finite/countable/uncountable, then the cover is said to be a* finite/countable/uncountable cover.

We shall not consider the case of uncountable cover in this text, because uncountable additivity makes no sense.

Exercise 11. Every subset of $\mathbb{R}^n$ admits a countable covering!

If we wish to associate a unique positive number to $E \in 2^{\mathbb{R}^n}$, satisfying monotonicity and (finite/countable) sub-additivity, then the association must satisfy

$$\begin{aligned}
\mu(E) &\le \mu\left(\cup_{i \in I} R_i\right) \quad \text{(due to monotonicity)} \\
&\le \sum_{i \in I} \mu(R_i) \quad \text{(due to finite/countable sub-additivity)} \\
&= \sum_{i \in I} |R_i| \quad \text{(measure same as volume).}
\end{aligned}$$

The case when the index set I is strictly finite corresponds to Riemann integration which we wish to generalise. Thus, we let I to be a countable index set, henceforth. A brief note on the case when index set I is strictly finite is given in § 2.6.

Definition 2.2.2. *For a subset E of $\mathbb{R}^n$, we define its* Lebesgue outer measure[1] $\mu^\star(E)$ *as,*

$$\mu^\star(E) := \inf_{E \subseteq \cup_{i\in I} R_i} \sum_{i\in I} |R_i|,$$

the infimum being taken over all possible countable *coverings of E.*

The Lebesgue outer measure is a well-defined non-negative set function on the power set of $\mathbb{R}^n$, $2^{\mathbb{R}^n}$.

Exercise 12. The outer measure is unchanging if we restrict ourselves to open covering or closed covering, i.e., for every subset $E \subset \mathbb{R}^n$,

$$\mu^\star(E) = \inf_{E \subseteq \cup_{i\in I} S_i} \sum_{i\in I} |S_i|,$$

where the infimum is taken over all possible countable *closed* or *open* coverings $\{S_i\}$ of E.

Before we see some examples for calculating outer measures of subsets of $\mathbb{R}^n$, let us observe some immediate properties of outer measure following from definition.

Lemma 2.2.3. *The outer measure $\mu^\star$ has the following properties:*

(a) For every subset $E \subseteq \mathbb{R}^n$, $0 \le \mu^\star(E) \le +\infty$.

*(b) (**Translation Invariance**) For every $E \subset \mathbb{R}^n$, $\mu^\star(E + x) = \mu^\star(E)$ for all $x \in \mathbb{R}^n$.*

*(c) (**Monotone**) If $E \subset F$, then $\mu^\star(E) \le \mu^\star(F)$.*

*(d) (**Countable Sub-additivity**) If $E = \cup_{i=1}^\infty E_i$ then*

$$\mu^\star(E) \le \sum_{i=1}^\infty \mu^\star(E_i).$$

[1] Why we call it "outer" and superscript with a $\star$ will be clear in the next section

Proof. (a) The non-negativity of the outer measure is an obvious consequence of the non-negativity of $|\cdot|$.

(b) The invariance under translation is obvious too, by noting that for each covering $\{R_i\}$ of E or $E+x$, $\{R_i+x\}$ and $\{R_i-x\}$ is a covering of $E+x$ and E, respectively, and the volumes of the cell is invariant under translation.

(c) Monotonicity is obvious, by noting the fact that, the family of covering of F is a sub-family of the coverings of E. Thus, the infimum over the family of cover for E is smaller than the sub-family.

(d) If $\mu^\star(E_i) = +\infty$, for some i, then the result is trivially true. Thus, we assume that $\mu^\star(E_i) < +\infty$, for all i. By the definition of outer measure, for each $\varepsilon > 0$, there is a covering by cells $\{R^i_j\}_{j=1}^\infty$ for E_i such that

$$\sum_{j=1}^{\infty} |R^i_j| \leq \mu^\star(E_i) + \frac{\varepsilon}{2^i}.$$

Since $E = \cup_{i=1}^\infty E_i$, the family $\{R^i_j\}_{i,j}$ is a covering for E. Thus,

$$\mu^\star(E) \leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |R^i_j| \leq \sum_{i=1}^{\infty}\left(\mu^\star(E_i) + \frac{\varepsilon}{2^i}\right) = \sum_{i=1}^{\infty}\mu^\star(E_i) + \varepsilon.$$

Since choice of ε is arbitrary, we have the countable sub-additivity of $\mu^\star$. □

We have seen the properties of outer measure. Let us now compute the outer measure for some subsets of $\mathbb{R}^n$.

Example 2.1. Outer measure of the empty set is zero, $\mu^\star(\emptyset) = 0$. Every cell is a cover for the empty set. Thus, infimum over the volume of all cells is zero.

Example 2.2. The outer measure for a singleton set $\{x\}$ in $\mathbb{R}^n$ is zero. The same argument as for empty set holds except that now the infimum is taken over all cells containing x. Thus, for each $\varepsilon > 0$, one can find a cell R_ε such that $x \in R_\varepsilon$ and $|R_\varepsilon| \leq \varepsilon$. Therefore, $\mu^\star(\{x\}) \leq \varepsilon$ for all $\varepsilon > 0$ and hence $\mu^\star(\{x\}) = 0$.

Example 2.3. The outer measure of any countable subset E of $\mathbb{R}^n$ is zero. A countable set $E = \cup_{x\in E}\{x\}$, where the union is countable. Thus, by countable sub-additivity, $\mu^\star(E) \leq 0$ and hence $\mu^\star(E) = 0$. Let's highlight something interesting at this stage. Note that the set of all rationals, $\mathbb{Q}$, in $\mathbb{R}$ is countable. Hence $\mu^\star(\mathbb{Q}) = 0$. Also $\mathbb{Q}$ is dense in $\mathbb{R}$. Thus, we actually have a dense ('scattered') subset of $\mathbb{R}$ whose outer measure is zero ('small').

Example 2.4. The situation is even worse. The converse of above example is not true, i.e., we can have a uncountable set whose outer measure is zero. The outer measure of $\mathbb{R}^{n-1}$ (lower dimensional), for $n \geq 2$, as a subset of $\mathbb{R}^n$ is zero, i.e., $\mu^\star(\mathbb{R}^{n-1}) = 0$. Choose a cover $\{R_i\}$ of $\mathbb{R}^{n-1}$ in $\mathbb{R}^{n-1}$ such that $|R_i|_{n-1} = 1$. Then $E_i = R_i \times \left(\frac{-\varepsilon}{2^i}, \frac{\varepsilon}{2^i}\right)$ forms a cover for $\mathbb{R}^{n-1}$ in $\mathbb{R}^n$. Thus,

$$\begin{aligned}\mu^\star(\mathbb{R}^{n-1}) &\leq \sum_{i=1}^{\infty} |E_i| \\ &= 2\varepsilon \sum_{i=1}^{\infty} \frac{1}{2^i} = 2\varepsilon.\end{aligned}$$

Since the choice of $\varepsilon > 0$ could be as small as possible, we have $\mu^\star(\mathbb{R}^{n-1}) = 0$.

Example 2.5. Is there an uncountable subset of $\mathbb{R}$ whose outer measure zero? Consider the Cantor set C (cf. Appendix A) which is uncountable. Let us compute the outer measure of C. Recall that, for each i, C_i is disjoint union of 2^i closed intervals, each of whose length is 3^{-i}. Thus, $\mu^\star(C_i) = (2/3)^i$, for all i. By construction, $C \subset C_i$ and hence due to monotonicity $\mu^\star(C) \leq \mu^\star(C_i) = (2/3)^i$, for all i. But $(2/3)^i \stackrel{i\to\infty}{\longrightarrow} 0$. Thus, $\mu^\star(C) = 0$.

Example 2.6. A similar argument as above shows that the outer measure of the generalised Cantor set C (cf. Appendix A) is bounded above by

$$\mu^\star(C) \leq \lim_k 2^k a_1 a_2 \ldots a_k.$$

We shall, in fact, show that equality holds here using "continuity from above" of outer measure (cf. Example 2.12).

Example 2.7. If R is any cell of $\mathbb{R}^n$, then $\mu^\star(R) = |R|$. Since R is a cover by itself, we have from definition, $\mu^\star(R) \leq |R|$. It now remains to prove the reverse inequality. Let S be a closed cell. Let $\{R_i\}_1^\infty$ be an arbitrary covering of S. Choose an arbitrary $\varepsilon > 0$. For each i, we choose an open cell Q_i such that $R_i \subset Q_i$ and $|Q_i| < |R_i| + \varepsilon|R_i|$. Note that $\{Q_i\}$ is an *open* covering of

S. Since S is compact in $\mathbb{R}^n$ (closed and bounded), we can extract a finite sub-cover such that $S \subset \cup_{i=1}^k Q_i$. Therefore,

$$\begin{aligned} |S| &\leq \sum_{i=1}^{k} |Q_i| \quad \text{(cf. Exercise 6)} \\ &\leq (1+\varepsilon) \sum_{i=1}^{k} |R_i|. \end{aligned}$$

Since ε can be chosen as small as possible, we have

$$|S| \leq \sum_{i=1}^{k} |R_i| \leq \sum_{i=1}^{\infty} |R_i|.$$

Since $\{R_i\}$ was an arbitrary choice of cover for S, taking infimum, we get $|S| \leq \mu^\star(S)$. Thus, for a closed cell we have shown that $|S| = \mu^\star(S)$. Now, for the given cell R and $\varepsilon > 0$, one can always choose a closed cell $S \subset R$ such that $|R| < |S| + \varepsilon$. Thus,

$$|R| < |S| + \varepsilon = \mu^\star(S) + \varepsilon \leq \mu^\star(R) + \varepsilon \quad \text{(by monotonicity)}.$$

Since $\varepsilon > 0$ is arbitrary, $|R| < \mu^\star(R)$ and hence $|R| = \mu^\star(R)$, for any cell R.

Example 2.8. The outer measure of $\mathbb{R}^n$ is infinite, $\mu^\star(\mathbb{R}^n) = +\infty$. For any $M > 0$, every cell R of volume M is a subset of $\mathbb{R}^n$. Hence, by monotonicity of $\mu^\star$, $\mu^\star(R) \leq \mu^\star(\mathbb{R}^n)$. But $\mu^\star(R) = |R| = M$. Thus, $\mu^\star(\mathbb{R}^n) \geq M$ for all $M > 0$. Thus, $\mu^\star(\mathbb{R}^n) = +\infty$.

Exercise 13. Show that $\mathbb{R}$ is uncountable. Also show that the outer measure of the set of irrationals in $\mathbb{R}$ is $+\infty$.

Example 2.9. Does the outer measure of other basic subsets, such as balls (or spheres), polygons etc. coincide with their volume, which we know from geometry (calculus)? We shall postpone answering this, in a simple way, till we develop sufficient tools. However, we shall note the fact that for any non-empty open set $\Omega \subset \mathbb{R}^n$, its outer measure is non-zero, i.e., $\mu^\star(\Omega) > 0$. This is because for every non-empty open set Ω, one can always find a cell $R \subset \Omega$ such that $|R| > 0$. Thus, $\mu^\star(\Omega) \geq |R| > 0$.

Exercise 14. If $E \subset \mathbb{R}^n$ has a positive outer measure, $\mu^\star(E) > 0$, does there always exist a cell $R \subset E$ such that $|R| > 0$.

Exercise 15. Let $E \subset \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then $\mu^\star(T(E)) = |\det(T)|\mu^\star(E)$. Consequently, $\mu^\star$ has the following properties:

(i) (**Reflection**) $\mu^\star(E) = \mu^\star(-E)$ where $-E := \{-x \mid x \in E\}$.

(ii) (**Dilation**) for $\lambda > 0$, $\mu^\star(\lambda E) = \lambda^n \mu^\star(E)$ where $\lambda E := \{\lambda x \mid x \in E\}$.

(iii) $\mu^\star$ is invariant under rotations.

Recall the definition of outer measure, which was infimum over all covers made up of cells. By a cell, we meant a rectangle in $\mathbb{R}^n$ whose sides are parallel to the coordinate axes. As a consequence of above exercise, it turns out that the Lebesgue outer measure is *invariant* if we include rectangles whose sides are not parallel to the coordinate axes.

Theorem 2.2.4 (Outer Regularity)**.** *If $E \subset \mathbb{R}^n$ is a subset of $\mathbb{R}^n$, then $\mu^\star(E) = \inf_{\Omega \supset E} \mu^\star(\Omega)$, where Ω's are open sets containing E.*

Proof. By monotonicity, $\mu^\star(E) \le \mu^\star(\Omega)$ for all open sets Ω containing E. Thus, $\mu^\star(E) \le \inf \mu^\star(\Omega)$. Conversely, for each $\varepsilon > 0$, we can choose a cover of cells $\{R_i\}$ of E such that

$$\sum_{i=1}^{\infty} |R_i| \le \mu^\star(E) + \frac{\varepsilon}{2}.$$

For each R_i, choose an open cell $Q_i \supset R_i$ such that

$$|Q_i| \le |R_i| + \frac{\varepsilon}{2^{i+1}}.$$

Since each Q_i is open and countable union of open sets is open, $\Omega = \cup_{i=1}^{\infty} Q_i$ is open. Therefore, by sub-additivity,

$$\mu^\star(\Omega) \le \sum_{i=1}^{\infty} |Q_i| \le \sum_{i=1}^{\infty} |R_i| + \frac{\varepsilon}{2} \le \mu^\star(E) + \varepsilon.$$

Thus, we have equality. $\square$

Recall that arbitrary union (intersection) of open (closed) sets is open (closed) and finite intersection (union) of open (closed) sets is open (closed). This motivates us to define the notion of $\mathcal{G}_\delta$ and $\mathcal{F}_\sigma$ subset of $\mathbb{R}^n$.

Definition 2.2.5. *A subset E is said to be $\mathcal{G}_\delta$*[2] *if it is a countable intersection of open sets in $\mathbb{R}^n$. We say E is $\mathcal{F}_\sigma$*[3] *if it is a countable union of closed sets in $\mathbb{R}^n$.*

Corollary 2.2.6. *For every subset $E \subset \mathbb{R}^n$ there exists a $\mathcal{G}_\delta$ subset G of $\mathbb{R}^n$ such that $G \supset E$ and $\mu^\star(E) = \mu^\star(G)$.*

Proof. Using Theorem 2.2.4, we have that for every $k \in \mathbb{N}$, there is an open set $\Omega_k \supset E$ such that

$$\mu^\star(\Omega_k) \leq \mu^\star(E) + \frac{1}{k}.$$

Let $G := \cap_{k=1}^\infty \Omega_k$. Thus, G is a $\mathcal{G}_\delta$ set. G is non-empty because $E \subset G$ and hence $\mu^\star(E) \leq \mu^\star(G)$. For the reverse inequality, we note that $G \subset \Omega_k$, for all k, and by monotonicity

$$\mu^\star(G) \leq \mu^\star(\Omega_k) \leq \mu^\star(E) + \frac{1}{k} \quad \forall k.$$

Thus, $\mu^\star(G) = \mu^\star(E)$. □

Exercise 16 (Continuity from below for outer measure). Let $E_1, E_2, \ldots$ be subsets of $\mathbb{R}^n$ such that $E_1 \subseteq E_2 \subseteq \ldots$ and $E = \cup_{i=1}^\infty E_i$, then $\mu^\star(E) = \lim_{k\to\infty} \mu^\star(E_k)$.

Proof. By monotonicity of outer measure, we get $\lim_{k\to\infty} \mu^\star(E_k) \leq \mu^\star(E)$. It only remains to prove the reverse inequality. Let G_k be a $\mathcal{G}_\delta$ set such that $G_k \supset E_k$, for all k, and $\mu^\star(E_k) = \mu(G_k)$. The $\{G_k\}$ may not be an increasing sequence. Hence, we set $F_k := \cap_{j\geq k} G_j$ and $\{F_k\}$ is an increasing sequence of measurable sets. Also, $E_k \subset F_k$, since $E_k \subset G_j$ for all $j \geq k$. Moreover, since $E_k \subset F_k \subset G_k$, we have $\mu^\star(E_k) \leq \mu(F_k) \leq \mu(G_k)$ and hence $\mu^\star(E_k) = \mu(F_k) = \mu(G_k)$. Since $E \subset \cup_{k=1}^\infty F_k$

$$\mu^\star(E) \leq \mu(\cup_{k=1}^\infty F_k) = \lim_{k\to\infty} \mu(F_k) = \lim_{k\to\infty} \mu^\star(E_k).$$

□

[2]terminology comes from German word "Gebiete" and "Durschnitt" meaning territory and average or mean, respectively

[3]terminology comes from French word "fermé" and "somme" meaning closed and sum, respectively

We have checked all the desired properties of $\mu^\star$, the outer measure, except the countable additivity. If $\mu^\star$ satisfies countable additivity, then we are done with our search for a 'measure' generalising the notion of volume. However, unfortunately, it turns out that $\mu^\star$ is not countably additive. In fact, in retrospect, this was the reason for naming $\mu^\star$ as "outer measure" instead of calling it "measure". Before we show that $\mu^\star$ is not countably additive, we show a property close to additivity.

Proposition 2.2.7. *If E and F are subsets of $\mathbb{R}^n$ such that $d(E,F) > 0$, then $\mu^\star(E \cup F) = \mu^\star(E) + \mu^\star(F)$.*

Proof. By the countable sub-additivity of $\mu^\star$, we have $\mu^\star(E \cup F) \leq \mu^\star(E) + \mu^\star(F)$. Once we show the reverse inequality, we are done. For each $\varepsilon > 0$, we can choose a covering $\{R_i\}$ of $E \cup F$ such that

$$\sum_{i=1}^{\infty} |R_i| \leq \mu^\star(E \cup F) + \varepsilon.$$

The family of cells $\{R_i\}$ can be categorised in to three groups: those intersecting only E, those intersecting only F and those intersecting both E and F. Note that the third category, cells intersecting both E and F, should have diameter bigger than $d(E,F)$. Thus, by subdividing these cells to have diameter less than $d(E,F)$, we can have the family of open cover to consist of only those cells which either intersect with E or F. Let $I_1 = \{i : R_i \cap E \neq \emptyset\}$ and $I_2 = \{i : R_i \cap F \neq \emptyset\}$. Due to our subdivision, we have $I_1 \cap I_2 = \emptyset$. Thus, $\{R_i\}$ for $i \in I_1$ is an open cover for E and $\{R_i\}$ for $i \in I_2$ is an open cover for F. Thus,

$$\mu^\star(E) + \mu^\star(F) \leq \sum_{i \in I_1} |R_i| + \sum_{i \in I_2} |R_i| \leq \sum_{i=1}^{\infty} |R_i| \leq \mu^\star(E \cup F) + \varepsilon.$$

By the arbitrariness of ε, we have the reverse inequality. $\square$

Proposition 2.2.8. *If a subset $E \subset \mathbb{R}^n$ is a countable union of almost disjoint closed cells, i.e., $E = \cup_{i=1}^{\infty} R_i$ then*

$$\mu^\star(E) = \sum_{i=1}^{\infty} |R_i|.$$

Proof. By countable sub-additivity, we already have $\mu^\star(E) \leq \sum_{i=1}^\infty |R_i|$. It only remains to prove the reverse inequality. For each $\varepsilon > 0$, let $Q_i \subset R_i$ be a cell such that $|R_i| \leq |Q_i| + \varepsilon/2^i$. By construction, the cells Q_i are pairwise disjoint and $d(Q_i, Q_j) > 0$, for all $i \neq j$. Applying Proposition 2.2.7 finite number times, we have

$$\mu^\star\left(\cup_{i=1}^k Q_i\right) = \sum_{i=1}^k |Q_i| \quad \text{for each } k \in \mathbb{N}.$$

Since $\cup_{i=1}^k Q_i \subset E$, by monotonicity, we have

$$\mu^\star(E) \geq \mu^\star\left(\cup_{i=1}^k Q_i\right) = \sum_{i=1}^k |Q_i| \geq \sum_{i=1}^k \left(|R_i| - \varepsilon/2^i\right)$$

By letting $k \to \infty$, we deduce

$$\sum_{i=1}^\infty |R_i| \leq \mu^\star(E) + \varepsilon.$$

Since ε can be made arbitrarily small, we have equality. □

A consequence above proposition and Theorem 2.1.4 is that for an open set Ω,

$$\mu^\star(\Omega) = \sum_{i=1}^\infty |R_i|$$

irrespective of the choice of the almost disjoint closed cells whose union is Ω.

We now show that $\mu^\star$ is not countably additive. In fact, it is not even finitely additive. One should observe that finite additivity of $\mu^\star$ is quite different from the result proved in Proposition 2.2.7. To show the non-additivity (finite) of $\mu^\star$, we need to find two disjoint sets, sum of whose outer measure is not the same as the outer measure of their union[4].

Proposition 2.2.9. *There exists a countable family $\{N_i\}_1^\infty$ of disjoint subsets of $\mathbb{R}^n$ such that*

$$\mu^\star\left(\cup_{i=1}^\infty N_i\right) \neq \sum_{i=1}^\infty \mu^\star(N_i).$$

[4]This construction is related to the Banach-Tarski paradox which states that one can partition the unit ball in $\mathbb{R}^3$ in to a finite number of pieces which can be reassembled (after rotation and translation) to form two complete unit balls!

Proof. Consider the unit cube $[0,1]^n$ in $\mathbb{R}^n$. We define an equivalence relation $\sim$ (cf. Appendix **??**) on $[0,1]^n$ as, $x \sim y$ whenever $x-y \in \mathbb{Q}^n$, i.e. we consider the quotient space $[0,1]^n/\mathbb{Q}^n$. This equivalence relation will partition the cube $[0,1]^n$ in to disjoint equivalence classes $\mathcal{E}_\alpha$, $[0,1]^n = \cup_\alpha \mathcal{E}_\alpha$. Now, let N be the subset of $[0,1]^n$ which is formed by picking[5] exactly one element from each equivalence $\mathcal{E}_\alpha$. Since $\mathbb{Q}^n$ is countable, let $\{r_i\}_1^\infty$ be the enumeration of all elements of $\mathbb{Q}^n$. Let $N_i := N + r_i$. We first show that N_i's are all pairwise disjoint. Suppose $N_i \cap N_j \neq \emptyset$, then there exist $x_\alpha, x_\beta \in N$ such that $x_\alpha + r_i = x_\beta + r_j$. Hence $x_\alpha - x_\beta = r_j - r_i \in \mathbb{Q}^n$. This implies that $x_\alpha \sim x_\beta$ which contradicts that fact that N contains exactly one representative from each equivalence class. Thus, N_i's are all disjoint. We now show that $\cup_{i=1}^\infty N_i = \mathbb{R}^n$. It is obvious that $\cup_{i=1}^\infty N_i \subset \mathbb{R}^n$. To show the reverse inclusion, we consider $x \in \mathbb{R}^n$. Then there is a $r_k \in \mathbb{Q}^n$ such that $x \in [0,1]^n + r_k$. Hence $x - r_k \in [0,1]^n$. Thus x belongs to some equivalence class, i.e., there is a $x_\alpha \in N$ such that $x - r_k \sim x_\alpha$. Therefore, $x \in N_k$. Hence $\cup_{i=1}^\infty N_i = \mathbb{R}^n$. Using the sub-additivity and translation-invariance of $\mu^\star$, we have

$$+\infty = \mu^\star(\mathbb{R}^n) \le \sum_{i=1}^\infty \mu^\star(N_i) = \sum_{i=1}^\infty \mu^\star(N).$$

Thus, $\mu^\star(N) \neq 0$, i.e., $\mu^\star(N) > 0$.

Let $J := \{i \mid r_i \in \mathbb{Q}^n \cap [0,1]^n\}$. Note that J is countable. Now, consider the sub-collection $\{N_j\}_{j\in J}$ and let $F = \cup_{j\in J} N_j$. Being a sub-collection of N_i's, N_j's are disjoint. By construction, each $N_j \subseteq [0,2]^n$, and hence $F \subset [0,2]^n$. By monotonicity and volume of cell, $\mu^\star(F) \le 2^n$. If countable-additivity holds true for the sub-collection N_j, then

$$\mu^\star(F) = \sum_{j\in J} \mu^\star(N_j) = \sum_{j\in J} \mu^\star(N).$$

$\mu^\star(F)$ is either zero or $+\infty$. But we have already shown that $\mu^\star(N) \neq 0$ and hence $\mu^\star(F) \neq 0$. Thus, $\mu^\star(F) = +\infty$, which contradicts $\mu^\star(F) \le 2^n$. Thus, countable-additivity for N_j cannot hold true. □

Remark 2.2.10. The clever construction of the set N in the above proof is due to Giuseppe Vitali. Thus, the set constructed above is called the *Vitali* set. There are more than one Vitali set. Each choice of representative from the equivalence class yields a different Vitali set.

[5] Possible due to Axiom of Choice

Corollary 2.2.11. *There exists a finite family $\{N_i\}_1^k$ of disjoint subsets of $\mathbb{R}^n$ such that*

$$\mu^\star\left(\cup_{i=1}^k N_i\right) \neq \sum_{i=1}^k \mu^\star(N_i).$$

Proof. The proof is ditto till proving the fact that $\mu^\star(N) > 0$. Now, it is always possible to choose a $k \in \mathbb{N}$ such that $k\mu^\star(N) > 2^n$. Then, we pick exactly k elements from the set J and for F to be the finite (k) union of $\{N_j\}_1^k$. Now, arguing as above assuming finite-additivity will contradict the fact that $k\mu^\star(N) > 2^n$. □

The fact that the outer measure $\mu^\star$ is not countably (even finitely) additive[6] is a bad news and leaves our job of generalising the notion of volume, for all subsets of $\mathbb{R}^n$, incomplete.

2.2.1 Abstract Set-up

Let X be any non-empty set. A set function $\mu^\star : 2^X \to [0, \infty]$ is called a *outer measure* on X if

(i) $\mu^\star(\emptyset) = 0$

(ii) If $E \subset \cup_{i=1}^\infty E_i$ then $\mu^\star(E) \leq \sum_{i=1}^\infty \mu^\star(E_i)$.

An outer measure $\mu^\star$ is said to be *finite* if $\mu^\star(X) < \infty$. An outer measure $\mu^\star$ is said to be *σ-finite*, if there exists subsets $\{E_i\}$ of X such that $\mu^\star(E_i) < +\infty$, for all i, and $X = \cup_{i=1}^\infty E_i$.

Exercise 17. Show that there exists subsets $\{E_i\}$ such that $\mu^\star(E_i) < +\infty$ and $\mathbb{R}^n = \cup_{i=1}^\infty E_i$, i.e. $\mathbb{R}^n$ is σ-finite with respect to the Lebesgue outer measure on $\mathbb{R}^n$.

2.3 Measurable Sets

The countable non-additivity of the outer measure, $\mu^\star$, has left our job incomplete. The next possible attempt is to consider only those subsets of $\mathbb{R}^n$ for which $\mu^\star$ is countably additive, i.e. we no longer work in the power set

[6]However, it is possible to get a finitely additive set function on $2^{\mathbb{R}^n}$ which coincides with $\mu^\star$ on $\mathcal{L}(\mathbb{R}^n)$

$2^{\mathbb{R}^n}$ but relax ourselves to a subclass of $2^{\mathbb{R}^n}$ for which countable additivity holds. For the chosen sub-class of subsets, the notion of volume is generalised and hence we shall call the sub-class 'measurable' sets and the outer measure $\mu^\star$ restricted to the sub-class is called the 'measure' of the set.

However, the difficulty lies in identifying the sub-class? In order to characterize the class of measurable sets we observe as a consequence of Theorem 2.2.4 that, for any subset $E \subset \mathbb{R}^n$ and for each $\varepsilon > 0$, there is an open set $\Omega \supset E$ such that $\mu^\star(\Omega) \leq \mu^\star(E) + \varepsilon$ or $\mu^\star(\Omega) - \mu^\star(E) \leq \varepsilon$. Since $\Omega = E \cup \Omega \setminus E$ is a disjoint union, by demanding countable additivity, we expect to have $\mu^\star(\Omega) - \mu^\star(E) \geq \mu^\star(\Omega \setminus E)$[7]. Thus, we need to choose those subsets of $\mathbb{R}^n$ for which $\mu^\star(\Omega) - \mu^\star(E) \geq \mu^\star(\Omega \setminus E)$.

Definition 2.3.1. *We say a subset $E \subset \mathbb{R}^n$ is* measurable (Lebesgue)*, if for any $\varepsilon > 0$ there exists an open set $\Omega \supset E$, containing E, such that $\mu^\star(\Omega \setminus E) \leq \varepsilon$. Further, we define the* measure (Lebesgue)*, μ, of E as $\mu(E) = \mu^\star(E)$.*

Let $\mathcal{L}(\mathbb{R}^n)$ denote the class of all subsets of $\mathbb{R}^n$ which are Lebesgue measurable. Thus, $\mathcal{L}(\mathbb{R}^n) \subset 2^{\mathbb{R}^n}$. The domain of outer measure $\mu^\star$, is $2^{\mathbb{R}^n}$, whereas the domain for the Lebesgue measure, μ, is $\mathcal{L}(\mathbb{R}^n)$. The inclusion $\mathcal{L}(\mathbb{R}^n) \subset 2^{\mathbb{R}^n}$ is proper (cf. Exercise 22).

By definition, the Lebesgue measure, μ will inherit all the properties of the outer measure $\mu^\star$. We shall see some examples of Lebesgue measurable subsets of $\mathbb{R}^n$.

Example 2.10. It is easy to see that every open set in $\mathbb{R}^n$ belongs to $\mathcal{L}(\mathbb{R}^n)$. Thus, $\emptyset, \mathbb{R}^n$, open cells etc. are all in $\mathcal{L}(\mathbb{R}^n)$.

Example 2.11. Every subset E of $\mathbb{R}^n$ such that $\mu^\star(E) = 0$ is in $\mathcal{L}(\mathbb{R}^n)$. By Theorem 2.2.4, for any $\varepsilon > 0$, there is an open set $\Omega \supseteq E$ such that $\mu^\star(\Omega) \leq \mu^\star(E) + \varepsilon = \varepsilon$. Since $\Omega \setminus E \subseteq \Omega$, by monotonicity of $\mu^\star$, we have $\mu^\star(\Omega \setminus E) \leq \varepsilon$. Thus, $E \in \mathcal{L}(\mathbb{R}^n)$. As a consequence, all singletons, finite set, countable sets, Cantor set C in $\mathbb{R}$, $\mathbb{R}^{n-1} \subset \mathbb{R}^n$ etc. are all in $\mathcal{L}(\mathbb{R}^n)$.

Theorem 2.3.2. *If $\{E_i\}_1^\infty$ is a countable family in $\mathcal{L}(\mathbb{R}^n)$, then $E := \cup_{i=1}^\infty E_i$ is in $\mathcal{L}(\mathbb{R}^n)$, i.e., countable union of measurable sets is measurable.*

Proof. Since each E_i is measurable, for any $\varepsilon > 0$, there is an open set $\Omega_i \supset E_i$ such that

$$\mu^\star(\Omega_i \setminus E_i) \leq \frac{\varepsilon}{2^i}.$$

[7]The other inequality being true due to sub-additivity

Let $\Omega = \cup_{i=1}^{\infty}\Omega_i$, then $E \subset \Omega$ and Ω is open. But $\Omega \setminus E \subset \cup_{i=1}^{\infty}(\Omega_i \setminus E_i)$. By monotonicity and sub-additivity of $\mu^\star$,

$$\mu^\star(\Omega \setminus E) \leq \mu^\star\left(\cup_{i=1}^{\infty}\Omega_i \setminus E_i\right) \leq \sum_{i=1}^{\infty} \mu^\star(\Omega_i \setminus E_i) \leq \varepsilon.$$

Thus, E is measurable. □

Definition 2.3.3. *We say $E \subset \mathbb{R}^n$ is bounded if there is a cell $R \subset \mathbb{R}^n$ of finite volume such that $E \subset R$.*

Exercise 18. If E is bounded then show that $\mu^\star(E) < +\infty$. Also, give an example of a set with $\mu^\star(E) < +\infty$ but E is unbounded.

Proposition 2.3.4. *Compact subsets of $\mathbb{R}^n$ are measurable.*

Proof. Let F be a compact subset of $\mathbb{R}^n$. Thus, $\mu^\star(F) < +\infty$. By Theorem 2.2.4, we have, for each $\varepsilon > 0$, an open subset $\Omega \supset F$ such that

$$\mu^\star(\Omega) \leq \mu^\star(F) + \varepsilon.$$

If we show $\mu^\star(\Omega \setminus F) \leq \varepsilon$, we are done. Observe that $\Omega \setminus F$ is an open set (since F is closed) and hence, by Theorem 2.1.4, there exists almost disjoint closed cells R_i such that

$$\Omega \setminus F = \cup_{i=1}^{\infty} R_i.$$

For a fixed $k \in \mathbb{N}$, consider the finite union of the closed cells $K := \cup_{i=1}^{k} R_i$. Note that K is compact. Also $K \cap F = \emptyset$ and thus, by Lemma **??**, $d(F, K) > 0$. But $K \cup F \subset \Omega$. Thus,

$$\begin{aligned}
\mu^\star(\Omega) &\geq \mu^\star(K \cup F) \quad \text{(Monotonicity)} \\
&= \mu^\star(K) + \mu^\star(F) \quad \text{(By Proposition 2.2.7)} \\
&= \mu^\star(\cup_{i=1}^{k} R_i) + \mu^\star(F) \\
&= \sum_{i=1}^{k} |R_i| + \mu^\star(F) \quad \text{(By Proposition 2.2.8)}.
\end{aligned}$$

Hence, $\sum_{i=1}^{k} |R_i| \leq \mu^\star(\Omega) - \mu^\star(F) \leq \varepsilon$. Since this is true for every $k \in \mathbb{N}$, by taking limit we have $\sum_{i=1}^{\infty} |R_i| \leq \varepsilon$. Using Proposition 2.2.8 again, we have

$$\mu^\star(\Omega \setminus F) = \sum_{i=1}^{\infty} |R_i| \leq \varepsilon.$$

Thus, F is measurable. □

Corollary 2.3.5. *Closed subsets of $\mathbb{R}^n$ are measurable. Consequently, any $\mathcal{F}_\sigma$ set is measurable.*

Proof. Let Γ be a closed subset of $\mathbb{R}^n$. If Γ were bounded, we know it is compact and hence measurable from the above proposition. We need to check only for unbounded set Γ. Let $F_i = \Gamma \cap \overline{B}_i(0)$, for $i = 1, 2, \ldots$. Note that each F_i is compact and $\Gamma = \cup_{i=1}^\infty F_i$. Since each F_i is measurable, using Theorem 2.3.2, we deduce that Γ is measurable.

Any $\mathcal{F}_\sigma$ set is a countable union of closed sets and hence is measurable. $\square$

Theorem 2.3.6. *If $E \in \mathcal{L}(\mathbb{R}^n)$ then $E^c \in \mathcal{L}(\mathbb{R}^n)$.*

Proof. Since $E \in \mathcal{L}(\mathbb{R}^n)$, for each $k \in \mathbb{N}$, there exists an open set $\Omega_k \supset E$ such that $\mu^\star(\Omega_k \setminus E) \leq 1/k$. Since Ω_k^c is closed, it is in $\mathcal{L}(\mathbb{R}^n)$. Set $F := \cup_{k=1}^\infty \Omega_k^c$. Note that F is a $\mathcal{F}_\sigma$ set and hence measurable. Since $\Omega_k^c \subset E^c$ for every k, we have $F \subset E^c$. Also $E^c \setminus F \subset \Omega_k \setminus E$, for all $k \in \mathbb{N}$. By monotonicity, $\mu^\star(E^c \setminus F) \leq 1/k$, for all $k \in \mathbb{N}$. Therefore, $\mu^\star(E^c \setminus F) = 0$ and hence is measurable. Now $E^c = (E^c \setminus F) \cup F$ is a union of two measurable sets and hence is measurable. $\square$

Corollary 2.3.7. *If $\{E_i\}_1^\infty$ is a countable family in $\mathcal{L}(\mathbb{R}^n)$, then $E := \cap_{i=1}^\infty E_i$ is in $\mathcal{L}(\mathbb{R}^n)$, i.e., countable intersection of measurable sets is measurable. Consequently, any $\mathcal{G}_\delta$ set is measurable.*

Proof. Note that $E = \cap_{i=1}^\infty E_i = (\cup_{i=1}^\infty E_i^c)^c$. $\square$

Exercise 19. If $E, F \in \mathcal{L}(\mathbb{R}^n)$ then show that $E \setminus F \in \mathcal{L}(\mathbb{R}^n)$.

Theorem 2.3.8 (Borel Regularity)**.** *For any subset $E \subset \mathbb{R}^n$, the following are equivalent:*

(i) $E \in \mathcal{L}(\mathbb{R}^n)$.

(ii) *For each $\varepsilon > 0$, there is an open set $\Omega \supset E$ such that $\mu(\Omega \setminus E) \leq \varepsilon$.*

(iii) *(**Inner regularity**) For each $\varepsilon > 0$, there is a closed set $\Gamma \subset E$ such that $\mu(E \setminus \Gamma) \leq \varepsilon$.*

(iv) *There exists a $\mathcal{F}_\sigma$ subset F of $\mathbb{R}^n$ such that $F \subset E$ and $\mu^\star(E \setminus F) = 0$.*

Proof. (i) implies (ii). Let E be measurable. Thus, for each $\varepsilon > 0$, there is an open set $\Omega \supset E$ such that

$$\mu^\star(\Omega \setminus E) \leq \varepsilon.$$

Since $\Omega \setminus E = \Omega \cap E^c$, it is measurable (intersection of measurable sets). Thus, $\mu^\star(\Omega \setminus E) = \mu(\Omega \setminus E) \leq \varepsilon$.

(i) and (ii) implies (iii). Let E be measurable and, thus, E^c is measurable. Applying (ii) to E^c we have, for each $\varepsilon > 0$, there is an open set $\Omega \supset E^c$ such that

$$\mu^\star(\Omega \setminus E^c) \leq \varepsilon.$$

Set $\Gamma := \Omega^c$. Then $\Gamma \subset E$. Note that $E \setminus \Gamma = \Omega \setminus E^c$. Hence $\mu(E \setminus \Gamma) \leq \varepsilon$.

(iii) implies (iv). Using (iii), we have that for every $k \in \mathbb{N}$, there is a closed set $\Gamma_k \subset E$ such that

$$\mu(E \setminus \Gamma_k) \leq \frac{1}{k}.$$

Let $F := \cup_{k=1}^\infty \Gamma_k$. Thus, F is a $\mathcal{F}_\sigma$ set and $F \subset E$. Note that $E \setminus F \subset E \setminus \Gamma_k$, for each k. Hence, by monotonicity, $\mu(E \setminus F) \leq \mu(E \setminus \Gamma_k) \leq 1/k$ for all k. Thus, $\mu(E \setminus F) = 0$.

(iv) implies (i). Assume (iv). Since F is a $\mathcal{F}_\sigma$ set, it is measurable. And since $E \setminus F$ has outer measure zero, we have $E \setminus F$ is measurable. Now, since $E = F \cup (E \setminus F)$, it is measurable. □

Exercise 20. If $E \subset \mathbb{R}^m$ and $F \subset \mathbb{R}^n$ are measurable subsets, then $E \times F \subset \mathbb{R}^{m+n}$ is measurable and $\mu(E \times F) = \mu(E)\mu(F)$. (Hint: do by cases, for open sets, $\mathcal{G}_\delta$ sets, measure zero sets and then arbitrary sets.)

We have reached the climax of our search for a notion that generalises volume of cells. We shall now show that for the collection of Lebesgue measurable sets $\mathcal{L}(\mathbb{R}^n)$, countable additivity is true. We proved in Proposition 2.2.7 and 2.2.8 special cases of countable additivity.

Theorem 2.3.9 (Countable Additivity)**.** *If* $\{E_i\}_1^\infty$ *are collection of disjoint measurable sets and* $E := \cup_{i=1}^\infty E_i$*, then*

$$\mu(E) = \sum_{i=1}^\infty \mu(E_i).$$

Proof. Note that E is measurable, since it is countable union of measurable sets. Therefore, by sub-additivity, we have

$$\mu(E) \leq \sum_{i=1}^{\infty} \mu(E_i).$$

We need to prove the reverse inequality. Let us first assume that each E_i is bounded. Then, by inner regularity, there is a closed set $F_i \subset E_i$ such that $\mu^\star(E_i \setminus F_i) \leq \varepsilon/2^i$. By sub-additivity, $\mu^\star(E_i) \leq \mu^\star(F_i) + \varepsilon/2^i$. Each F_i is also pairwise disjoint, bounded and hence compact. Thus, by Lemma **??**, $d(F_i, F_j) > 0$ for all $i \neq j$. Therefore, using Proposition 2.2.7, we have for every $k \in \mathbb{N}$,

$$\mu\left(\cup_{i=1}^{k} F_i\right) = \sum_{i=1}^{k} \mu(F_i).$$

Since $\cup_{i=1}^{k} F_i \subset E$, by monotonicity, we have

$$\mu(E) \geq \mu\left(\cup_{i=1}^{k} F_i\right) = \sum_{i=1}^{k} \mu(F_i) \geq \sum_{i=1}^{k} \left(\mu(E_i) - \frac{\varepsilon}{2^i}\right).$$

Since the above inequality is true for every k and arbitrarily small ε, we get $\mu(E) \geq \sum_{i=1}^{\infty} \mu(E_i)$ and hence equality.

Let E_i be unbounded for some or all i. Consider a sequence of cells $\{R_j\}_1^\infty$ such that $R_j \subset R_{j+1}$, for all $j = 1, 2, \ldots$, and $\cup_{j=1}^{\infty} R_j = \mathbb{R}^n$. Set $Q_1 := R_1$ and $Q_j := R_j \setminus R_{j-1}$ for all $j \geq 2$. Consider the measurable subsets $Ei, j := E_i \cap Q_j$. Note the each $E_{i,j}$ is pairwise disjoint and are each bounded. Observe that $E_i = \cup_{j=1}^{\infty} E_{i,j}$ and is a disjoint union. Therefore, $\mu(E_i) = \sum_{j=1}^{\infty} \mu(E_{i,j})$. Also, $E = \cup_i \cup_j E_{i,j}$ and is a disjoint union. Hence,

$$\mu(E) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu(E_{i,j}) = \sum_{i=1}^{\infty} \mu(E_i).$$

□

Exercise 21. If $E \subseteq F$ and $\mu(E) < +\infty$, then show that $\mu(F \setminus E) = \mu(F) - \mu(E)$.

Corollary 2.3.10. *Let $E_1, E_2, \ldots$ be measurable subsets of $\mathbb{R}^n$.*

(i) *(**Continuity from below**) If $E_1 \subseteq E_2 \subseteq \dots$ and $E = \cup_{i=1}^{\infty} E_i$, then $\mu(E) = \lim_{k\to\infty} \mu(E_k)$.*

(ii) *(**Continuity from above**) If $E_1 \supseteq E_2 \supseteq \dots$, $\mu(E_i) < +\infty$, for some i, and $E = \cap_{i=1}^{\infty} E_i$, then $\mu(E) = \lim_{k\to\infty} \mu(E_k)$.*

Proof. (i) Let $F_1 := E_1$ and $F_i := E_i \setminus E_{i-1}$, for $i \geq 2$. By construction, F_i are measurable, disjoint and $E = \cup_{i=1}^{\infty} F_i$. Hence, by countable additivity,

$$\mu(E) = \sum_{i=1}^{\infty} \mu(F_i) = \lim_{k\to\infty} \sum_{i=1}^{k} \mu(F_i) = \lim_{k\to\infty} \mu(\cup_{i=1}^{k} F_i) = \lim_{k\to\infty} \mu(E_k).$$

(ii) Without loss of generality, we assume that $\mu(E_1) < +\infty$. Set $F_i = E_i \setminus E_{i+1}$, for each $i \geq 1$. Note that $E_1 = E \cup (\cup_{i=1}^{\infty} F_i)$ is a disjoint union of measurable sets. Hence,

$$\begin{aligned}
\mu(E_1) &= \mu(E) + \sum_{i=1}^{\infty} \mu(F_i) = \mu(E) + \sum_{i=1}^{\infty} (\mu(E_i) - \mu(E_{i+1}) \\
&= \mu(E) + \lim_{k\to\infty} \sum_{i=1}^{k-1} (\mu(E_i) - \mu(E_{i+1})) \\
&= \mu(E) + \mu(E_1) - \lim_{k\to\infty} \mu(E_k) \\
\lim_{k\to\infty} \mu(E_k) &= \mu(E).
\end{aligned}$$

□

Remark 2.3.11. Observe that for continuity from above, the assumption $\mu(E_i) < +\infty$ is very crucial. For instance, consider $E_i = (i, \infty) \subset \mathbb{R}$. Note that each $\mu(E_i) = +\infty$ but $\mu(E) = 0$.

Example 2.12. Recall that in Example 2.6 we proved an inequality regarding the generalised Cantor set C. We now have enough tools to show the equality. Note that each measurable C_k's satisfy the hypothesis of continuity from above and hence C is measurable and

$$\mu(C) = \lim_{k} 2^k a_1 a_2 \dots a_k.$$

In view of this example and Proposition A.0.6, we have generalised Cantor set whose outer measure is positive.

Recall that we showed the inner regularity of μ in Theorem 2.3.8. We can, in fact, better this for sets with finite measure.

Corollary 2.3.12. *If $\mu(E) < +\infty$ then there exists a compact set $K \subset E$ such that $\mu(E \setminus K) \leq \varepsilon$.*

Proof. We have, using (iii) of Theorem 2.3.8, that a closed set $\Gamma \subset E$ such that $\mu(E \setminus \Gamma) \leq \varepsilon$. Let $K_i := \Gamma \cap \overline{B}_i(0)$ be a sequence of compact sets such that $\Gamma = \cup_{i=1}^{\infty} K_i$ and $K_1 \subset K_2, \ldots$. Therefore, $E \setminus K_1 \supset E \setminus K_2 \supset \ldots$ and $E \setminus \Gamma = \cap_{i=1}^{\infty}(E \setminus K_i)$. Using, continuity from above and $\mu(E) < +\infty$, we get

$$\varepsilon \geq \mu(E \setminus \Gamma) = \lim_{i \to \infty} \mu(E \setminus K_i).$$

Thus, for i large enough, we have $\mu(E \setminus K_i) \leq \varepsilon$. $\square$

Theorem 2.3.13 (First Borel-Cantelli Lemma)**.** *If $\{E_i\}_1^{\infty} \subset \mathcal{L}(\mathbb{R}^n)$ be a countable collection of measurable subsets of $\mathbb{R}^n$ such that $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. Then $E := \cap_{k=1}^{\infty} \cup_{i=k}^{\infty} E_i$ has measure zero.*

Proof. Let $F_k := \cup_{i=k}^{\infty} E_i$. Note that $F_1 \supset F_2 \ldots$ and $E = \cap_{k=1}^{\infty} F_k$. Let $\sum_i \mu(E_i) = L$. By countable additivity, $\mu(F_1) \leq \sum_{i=1}^{\infty} \mu(E_i) = L < \infty$. By continuity from above, we have

$$\mu(E) = \lim_{k \to \infty} \mu(F_k) = \lim_{k \to \infty} \mu(\cup_{i=k}^{\infty} E_i) \leq \lim_{k \to \infty} \sum_{i=k}^{\infty} \mu(E_i) = \lim_{k \to \infty} (L - \sum_{i=1}^{k-1} \mu(E_i)).$$

Thus, $\mu(E) = 0$. $\square$

The set E in First Borel-Cantelli lemma is precisely the set of all $x \in \mathbb{R}^n$ such that $x \in E_i$ for infinitely many i. Let $x \in \mathbb{R}^n$ be such that $x \in E_i$ only for finitely many i. Arrange the indices i in increasing order, for which $x \in E_i$ and let K be the maximum of the indices. Then, $x \notin F_j$ for all $j \geq K + 1$ and hence not in E. Conversely, if $x \notin E$, then there exists a j such that $x \notin F_k$ for all $k \geq j$. Thus, either $x \in \cup_{i=1}^{j-1} E_i$ or $x \notin E_i$ for all i.

Exercise 22. Show that the Vitali set N constructed in Proposition 2.2.9 is not in $\mathcal{L}(\mathbb{R}^n)$. Thus, $\mathcal{L}(\mathbb{R}^n) \subset 2^{\mathbb{R}^n}$ is a strict inclusion.

Exercise 23. Consider the Vitali set N constructed in Proposition 2.2.9. Show that every measurable subset $E \subset N$ is of zero measure.

Proof. Let $E \subset N$ be a measurable set such that $\mu(E) > 0$, then for each $r_i \in \mathbb{Q} \cap [0,1]$, we set $E_i := E + r_i$ and $N_i := N + r_i$. Since N_i's are disjoint, E_i's are disjoint and are measurable. Since $\cup_{i=1}^\infty E_i \subset [0,2]$, we have

$$2 \geq \mu(\cup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty \mu(E) = +\infty.$$

A contradiction due to the assumption that $\mu(E) > 0$. Hence $\mu(E) = 0$. $\square$

Exercise 24. If $E \in \mathcal{L}(\mathbb{R}^n)$ such that $\mu(E) > 0$ then show that E has a non-measurable subset.

Proof. We first show that every measurable set $E \subset [0,1]$ such that $\mu(E) > 0$ has a non-measurable subset. Consider the non-measurable subset N of $[0,1]$. For each $r_i \in \mathbb{Q}$, we set $N_i := N + r_i$ and each of them are non-measurable. Also, we know that $\mathbb{R} = \cup_{i=1}^\infty N_i$. Let $E \subset [0,1]$ be a measurable subset such that $\mu(E) > 0$. Set $E_i := E \cap N_i$. Note that $\cup_{i=1}^\infty E_i = E \cap (\cup_{i=1}^\infty N_i) = E \cap \mathbb{R} = E$. If E_i were measurable, for each i, then being a subset of N_i, $\mu^\star(E_i) = 0$. Thus,

$$0 < \mu(E) = \mu^\star(\cup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \mu^\star(E_i) = 0,$$

a contradiction. Thus, our assumption that E_i are measurable is incorrect. Thus, E_i's are non-measurable subsets of E.

Now, let $E \subset \mathbb{R}$ be any measurable subset such that $\mu(E) > 0$. Note that $E = \cup_{i\in\mathbb{Z}} (E \cap [i, i+1))$, where $[i, i+1)$ are disjoint measurable subsets of $\mathbb{R}$. Hence,

$$0 < \mu(E) = \sum_{i\in\mathbb{Z}} \mu(E \cap [i, i+1)).$$

Thus, for some i, $\mu(E \cap [i, i+1)) > 0$. For this i, set $F := E \cap [i, i+1)$. Then $F - i \subset [0,1]$ which has positive measure and by earlier argument contains a non-measurable set M. Thus, $M + i \subset F \subset E$ is non-measurable. $\square$

Exercise 25. Construct an example of a continuous function that maps a measurable (Lebesgue) set to a non-measurable set.

Definition 2.3.14. *We say a subset $E \subset \mathbb{R}^n$ satisfies the* Carathéodory criterion *if*

$$\mu^\star(S) = \mu^\star(S \cap E) + \mu^\star(S \cap E^c) \quad \textit{for all subsets } S \subset \mathbb{R}^n.$$

Equivalently,

$$\mu^\star(S) = \mu^\star(S \cap E) + \mu^\star(S \setminus E) \quad \textit{for all subsets } S \subset \mathbb{R}^n.$$

Note that since $S = (S \cap E) \cup (S \cap E^c)$, by sub-additivity of $\mu^\star$, we always have

$$\mu^\star(S) \leq \mu^\star(S \cap E) + \mu^\star(S \cap E^c).$$

Thus, in order to check the Carathéodory criterion of a set E, it is enough to check the inequality

$$\mu^\star(S) \geq \mu^\star(S \cap E) + \mu^\star(S \cap E^c) \quad \forall S \subset \mathbb{R}^n \text{ and } \mu^\star(S) < +\infty.$$

Note that , intuitively, Carathéodory criterion classifies those sets that respect additivity.

The above criterion was given by Constantin Carathéodory for characterising the measurable sets. Some books also start with Carathéodory approach as the definition for measurability, since it is equivalent to our notion of measurability. Moreover, this has the advantage over our definition that it is topology independent and is a purely set-theoretic definition and will suit well in the abstract set-up.

Theorem 2.3.15. *$E \in \mathcal{L}(\mathbb{R}^n)$ if and only if E satisfies the Carathéodory criterion.*

Proof. Let $S \in \mathcal{L}(\mathbb{R}^n)$. Note that if $S \in \mathcal{L}(\mathbb{R}^n)$, then by countable additivity of μ we have the equality. Thus, it is enough to check the Carathéodory criterion of E with non-measurable sets S. Let S be any subset of $\mathbb{R}^n$ such that $\mu^\star(S) < +\infty$. We need to show that

$$\mu^\star(S) \geq \mu^\star(S \cap E) + \mu^\star(S \setminus E).$$

Corresponding to the set S, by Corollary 2.2.6, there is a $\mathcal{G}_\delta$ set (hence measurable) $G \supset S$ such that $\mu(G) = \mu^\star(S)$. By the countable additivity,

$$\mu^\star(S) = \mu(G) = \mu(G \cap E) + \mu(G \setminus E) \geq \mu^\star(S \cap E) + \mu^\star(S \setminus E)$$

where the last inequality is due to monotonicity. Hence one way implication is proved.

Conversely, let $E \subset \mathbb{R}^n$ satisfy the Carathéodory criterion. We need to show $E \in \mathcal{L}(\mathbb{R}^n)$. To avoid working with ∞, we assume E to be such that

$\mu^\star(E) < +\infty$. We know, by outer regularity, that for each $\varepsilon > 0$ there is an open set $\Omega \supset E$ such that $\mu^\star(\Omega) \leq \mu^\star(E) + \varepsilon$. But, by Carathéodory criterion, we have

$$\mu^\star(\Omega) = \mu^\star(\Omega \cap E) + \mu^\star(\Omega \setminus E) = \mu^\star(E) + \mu^\star(\Omega \setminus E).$$

Thus,

$$\mu^\star(\Omega \setminus E) = \mu^\star(\Omega) - \mu^\star(E) \leq \varepsilon.$$

Hence, E is measurable. It now only remains to prove for E such that $\mu^\star(E) = +\infty$. Let $E_i = E \cap \overline{B}_i(0)$, for $i = 1, 2, \ldots$. Note that each $\mu^\star(E_i) < +\infty$ is bounded and $E = \cup_{i=1}^\infty E_i$. Since each E_i is measurable, using Theorem 2.3.2, we deduce that E is measurable. □

2.3.1 Abstract Set-up

Let X be a non-empty set and 2^X is the collection of all subsets of X. We say a sub-collection $\mathcal{M} \subset 2^X$ of subsets of X to be a *σ-algebra* if

(i) $\emptyset \in \mathcal{M}$.

(ii) If $E \in \mathcal{M}$ then $E^c \in \mathcal{M}$ (closure under complementation).

(iii) If $\{E_i\} \subset \mathcal{M}$ then $\cup_i E_i \in \mathcal{M}$ (closure under countable union).

Exercise 26. Show that

(a) $2^{\mathbb{R}^n}$ is a σ-algebra.

(b) $\mathcal{L}(\mathbb{R}^n)$, the class of all Lebesgue measurable subsets of $\mathbb{R}^n$, forms a σ-algebra.

Let $\mathcal{M}$ be a σ-algebra. A set function $\mu : \mathcal{M} \to [0, \infty]$ is called a *measure* on X if

(i) $\mu(\emptyset) = 0$

(ii) If $E \subset \cup_{i=1}^\infty E_i$ then $\mu(E) \leq \sum_{i=1}^\infty \mu(E_i)$.

(iii) If $E = \cup_{i=1}^\infty E_i$ is a disjoint union then $\mu(E) = \sum_{i=1}^\infty \mu(E_i)$.

The triplet $(X, \mathcal{M}, \mu)$ is called a measure space.

Exercise 27. (a) Show that if $\{\mu_k\}$ is a sequence of measures on the same σ-algebra $\mathcal{M}$ then $\mu = \sum_k \mu_k$ is also a measure on $\mathcal{M}$.

(b) Show that Lebesgue measure is a measure on $\mathcal{L}(\mathbb{R}^n)$.

(c) Show that the cardinality of a set defines a measure on the σ-algebra 2^X. This is called the *counting measure*.

(d) Let X be infinite set. Define, for $E \subset X$,

$$\mu(E) = \begin{cases} 0 & E \text{ is countable} \\ +\infty & E \text{ otherwise.} \end{cases}$$

Show that μ is a measure on 2^X.

(e) Fix a $x \in X$. The Dirac measure at x, for $E \subset X$, is defined as

$$\delta_x(E) = \begin{cases} 1 & x \in E \\ 0 & \text{otherwise.} \end{cases}$$

Show that δ_x is a measure on 2^X, for each x.

Let X be a topological space and let $\tau(X)$ denote the collection of all open subsets of X. The smallest σ-algebra containing $\tau(X)$ is said to be the *Borel σ-algebra*. Every element of the Borel σ-algebra is called a *Borel Set*. Let $\mathcal{B}(\mathbb{R}^n)$ denote Borel σ-algebra of $\mathbb{R}^n$.

Exercise 28. Show that $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{L}(\mathbb{R}^n)$.

A measure space $(X, \mathcal{M}, \mu)$ is said to be *complete* if every subset of a set of measure zero belongs to $\mathcal{M}$.

Exercise 29 (Completeness). If $E \in \mathcal{L}(\mathbb{R}^n)$ such that $\mu(E) = 0$ then show that, for every $F \subset E$, $F \in \mathcal{L}(\mathbb{R}^n)$.

Exercise 30. The Lebesgue measure restricted to the Borel σ-algebra is not complete and its completion is $\mathcal{L}(\mathbb{R}^n)$.

A measure μ on a σ-algebra $\mathcal{M}$ is said to be a *probability measure* if $\mu(X) = 1$ and $0 \le \mu(E) \le 1$ for all $E \in \mathcal{M}$. The triplet $(X, \mathcal{M}, \mu)$ is called the *probability space*. Every element of X is called a *sample point*. Every element $E \in \mathcal{M}$ is called an *event* and $\mu(E)$ is the probability of the event E.

2.4 Measurable Functions

Recall that our aim was to develop the Lebesgue notion of integration for functions on $\mathbb{R}^n$. To do so, we need to classify those functions for which Lebesgue integration makes sense. We shall restrict our attention to real valued functions on $\mathbb{R}^n$. Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ denote the extended real line.

Definition 2.4.1. *We say a function f on $\mathbb{R}^n$ is* extended real valued *if it takes value on the extended real line $\overline{\mathbb{R}}$.*

Henceforth, we will confine ourselves to extended real valued functions unless stated otherwise. By a finite-valued function we will mean a function not taking $\pm\infty$.

Recall that we said Lebesgue integration is based on the idea of partitioning the range. As a simple case, consider the characteristic function of a set $E \subset \mathbb{R}^n$,

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E. \end{cases}$$

We expect that

$$\int_{\mathbb{R}^n} \chi_E(x) = \int_E = \mu^\star(E) = \mu(E).$$

But the last equality makes sense only when E is measurable. Thus, we expect to compute integrals of only those functions whose range when partitioned has the pre-image as a measurable subset of $\mathbb{R}^n$.

Definition 2.4.2. *We say a function $f : E \subset \mathbb{R}^n \to \overline{\mathbb{R}}$ is* measurable (Lebesgue) *if for all $\alpha \in \mathbb{R}$, the set*[8]

$$f^{-1}([-\infty, \alpha)) = \{f < \alpha\} := \{x \in E \mid f(x) < \alpha\}$$

is measurable (Lebesgue). We say f is Borel measurable *if $\{f < \alpha\}$ is a Borel set. A complex valued function $f : E \subset \mathbb{R}^n \to \mathbb{C}$ is said to be* measurable *if both its real and imaginary parts are measurable.*

Exercise 31. Every finite valued Borel measurable function is Lebesgue measurable.

[8]Usually, called as the sublevel set

Proposition 2.4.3. *If f is a finite-valued continuous function on $\mathbb{R}^n$ then f is Borel measurable. Consequently, every continuous function is Lebesgue measurable.*

Proof. Consider the open interval $I = (-\infty, \alpha)$ in $\mathbb{R}$. Since f is continuous, $f^{-1}(I) = \{f < \alpha\}$ is an open subset of $\mathbb{R}^n$ and hence is a Borel set. Thus, f is Borel measurable. □

The characteristic function χ_E is measurable but not continuous for a proper measurable subset of $\mathbb{R}^n$. Henceforth, by measurable function, we shall mean Lebesgue measurable function and by Borel function, we shall mean a Borel measurable function.

Exercise 32. (i) E is measurable iff χ_E is measurable.

(ii) f is measurable iff $\{f \leq \alpha\}$ is measurable for every $\alpha \in \mathbb{R}$.

(iii) f is measurable iff $\{f > \alpha\}$ is measurable for every $\alpha \in \mathbb{R}$.

(iv) If f is measurable then show that $-f$ is also measurable.

(v) Let f be finite-valued. f is measurable iff $\{\alpha < f < \beta\}$ is measurable for every $\alpha, \beta \in \mathbb{R}$.

Exercise 33. If f is measurable then

(i) f^k, is measurable for all integers $k \geq 1$.

(ii) $f + \lambda$ is measurable for a given constant $\lambda \in \mathbb{R}$.

(iii) λf is measurable for a given constant $\lambda \in \mathbb{R}$.

Exercise 34. If f, g are measurable and both finite-valued then both $f + g$ and fg are measurable.

Definition 2.4.4. *A property is said to hold* almost everywhere *(a.e.) if it holds except possibly on a set of measure zero.*

Consequently, we say a measurable function is finite a.e. if the set on which it takes $\pm\infty$ is of measure zero. All the "finite-valued" statements above can be replaced with "finite a.e.". Let $M(\mathbb{R}^n)$ denote the space of all finite a.e. measurable functions on $\mathbb{R}^n$. The class of functions $M(\mathbb{R}^n)$ forms a vector space over $\mathbb{R}$. Note that $M(\mathbb{R}^n)$ excludes those measurable function which takes values on extended line on a non-zero measure set.

Definition 2.4.5. *We say two functions f, g are equal a.e., $f = g$ a.e., if the set*

$$\{x \mid f(x) \neq g(x)\}.$$

is of measure zero.

Define the equivalence relation $f \sim g$ if $f = g$ a.e. on $M(\mathbb{R}^n)$. Thus, we have the quotient space $M(\mathbb{R}^n)/\sim$. However, as an abuse of notation, it has become standard to identify the quotient space $M(\mathbb{R}^n)/\sim$ with $M(\mathbb{R}^n)$. Henceforth, by $M(\mathbb{R}^n)$ we refer to the quotient space. Thus, whenever we say $A \subset M(\mathbb{R}^n)$, we usually mean the inclusion of the quotient spaces $A/\sim \subset M(\mathbb{R}^n)/\sim$. In other words, each equivalence class of $M(\mathbb{R}^n)$ has a representative from A.

Note that the support of a function $f \in M(\mathbb{R}^n)$ is defined as the closure of the set E,

$$E := \{x \mid f(x) \neq 0\}.$$

Thus, even though $\chi_{\mathbb{Q}} \sim 0$ are in the same equivalence class and represent the same element in $M(\mathbb{R}^n)/\sim$, the support of $\chi_{\mathbb{Q}}$ is $\mathbb{R}$ whereas the support of zero function is empty set. Therefore, whenever we say support of a function $f \in M(\mathbb{R}^n)$ has some property, we usually mean there is a representative in the equivalence class which satisfies the said properties.

Exercise 35 (Complete measure space). If f is measurable and $f = g$ a.e., then g is measurable.

Theorem 2.4.6. *Let f be finite a.e. on $\mathbb{R}^n$. The following are equivalent:*

(i) f is measurable.

(ii) $f^{-1}(\Omega)$ is a measurable set, for every open set Ω in $\mathbb{R}$.

(iii) $f^{-1}(\Gamma)$ is measurable for every closed set Γ in $\mathbb{R}$.

(iv) $f^{-1}(B)$ is measurable for every Borel set B in $\mathbb{R}$.

Proof. Without loss of generality we shall assume that f is finite valued on $\mathbb{R}^n$.

(i) implies (ii) Let Ω be an open subset of $\mathbb{R}$. Then $\Omega = \cup_{i=1}^{\infty}(a_1, b_i)$ where (a_i, b_i) are intervals which are pairwise disjoint. Observe that

$$\begin{aligned} f^{-1}(\Omega) &= \cup_{i=1}^{\infty} f^{-1}(a_i, b_i) \\ &= \cup_{i=1}^{\infty} \left(\{f > a_i\} \cap \{f < b_i\}\right). \end{aligned}$$

Since f is measurable, both $\{f > a_i\}$ and $\{f < b_i\}$ are measurable for all i. Since countable union and intersection of measurable sets are measurable, we have $f^{-1}(\Omega)$ is measurable.

(ii) implies (iii) $f^{-1}(\Gamma) = (f^{-1}(\Gamma^c))^c$. Since Γ^c is open, $f^{-1}(\Gamma^c)$ is measurable and complement of measurable sets are measurable.

(iii) implies (iv) Consider the collection

$$\mathcal{A} := \{F \subset \mathbb{R} \mid f^{-1}(F) \in \mathcal{L}(\mathbb{R}^n)\}.$$

We note that the collection $\mathcal{A}$ forms a σ-algebra. Firstly, $\emptyset \in \mathcal{A}$. Let $F \in \mathcal{A}$. $f^{-1}(F^c) = f^{-1}(\mathbb{R}) \setminus f^{-1}(F) = (f^{-1}(F))^c$. Also, $f^{-1}(\cup_{i=1}^\infty F_i) = \cup_{i=1}^\infty f^{-1}(F_i)$. By (iii), every closed subset $\Gamma \in \mathcal{A}$ and hence the Borel σ-algebra of $\mathbb{R}$ is included in $\mathcal{A}$.

(iv) implies (i) Let $f^{-1}(B)$ be measurable for every Borel set $B \subset \mathbb{R}$. In particular, $(-\infty, \alpha)$ is a Borel set of $\mathbb{R}$, for any $\alpha \in \mathbb{R}$, and hence $\{f < \alpha\}$ is measurable. Thus, f is measurable. $\square$

Exercise 36. If $f : \mathbb{R} \to \mathbb{R}$ is monotone then f is Borel.

The above exercise is useful in giving the example of a set which is Lebesgue measurable, but not Borel. Let $g : [0,1] \to [0,1]$ be the function defined as

$$g(y) = \inf_{x \in [0,1]} \{f_C(x) = y\}, \tag{2.4.1}$$

where f_C is the Cantor function (cf. Appendix A)

Exercise 37. Show that $g : [0,1] \to C$ is bijective and increasing. Consequently, g is Borel.

Example 2.13 (Example of a Measurable set which is not Borel). Let N be the non-measurable (Lebesgue) Vitali subset of $[0,1]$ (constructed in Proposition 2.2.9). Let $M := g(N)$ is a subset of C. Since $\mu^\star(C) = 0$, we have by monotonicity $\mu^\star(M) = 0$ and thus M is Lebesgue measurable (zero outer measure sets are measurable). If M were a Borel set then, by Borel measurability of g, $N = g^{-1}(M)$ is also Borel, a contradiction.

We provide another example using product of measures. Let N be the non-measurable (Lebesgue) Vitali subset of $\mathbb{R}$. Then $N \times \{0\} \subset \mathbb{R}^2$ has outer measure zero and hence is measurable subset of $\mathbb{R}^2$. If $N \times \{0\}$ were Borel set of $\mathbb{R}^2$ then N should be a Borel set of $\mathbb{R}$ (since it is a section of $N \times \{0\}$ with y coordinate fixed). But N is not Borel.

What about composition of measurable functions. Let $f : \mathbb{R}^n \to \mathbb{R}$ be measurable (Lebesgue) and $g : \mathbb{R} \to \mathbb{R}$ is also Lebesgue measurable. Then $g \circ f : \mathbb{R}^n \to \mathbb{R}$ need not be measurable because $g^{-1}(-\infty, \alpha)$ need not be a Borel set. However, by relaxing the condition on g, we may expect the composition to be measurable.

Proposition 2.4.7. *If f is measurable, finite a.e. on $\mathbb{R}^n$ and g is Borel measurable on $\mathbb{R}$ then $g \circ f$ is (Lebesgue) measurable. In particular, $g \circ f$ is measurable for continuous function g. Consequently, f^+, f^-, $|f|$ and $|f|^p$ for all $p > 0$ are all measurable. Also, for any two finite valued measurable functions f, g, $\max(f, g)$ and $\min(f, g)$ are measurable.*

Proof. Consider the set the interval $(-\infty, \alpha)$ in $\mathbb{R}$. By the Borel measurability of g, $\Omega := g^{-1}(-\infty, \alpha)$ is a Borel set of $\mathbb{R}$. Using the measurability of f, $(g \circ f)^{-1}(-\infty, \alpha) = f^{-1}(\Omega)$ is measurable. Prove the rest as an exercise. $\square$

Example 2.14. The reverse composition $f \circ g$, in general, is not (Lebesgue) measurable. Consider the Borel function $g : [0, 1] \to C$ given in (2.4.1). Let N be the non-measurable subset of $[0, 1]$. Set $E := g(N)$. Since $E \subset C$, E is Lebesgue measurable. Now, set $f := \chi_E$, which is measurable on $[0, 1]$. Observe that the composition $f \circ g = \chi_N$, is not Lebesgue measurable since N is a non-measurable set.

Exercise 38. We showed if f is measurable then $|f|$ is measurable. The converse is not true. Given an example of a non-measurable function f such that $|f|$ is measurable.

Proof. Let N denote the non-measurable subset (say, Vitali set) of $[0, 1]$. Define $f : [0, 1] \to \mathbb{R}$ as

$$f(x) = \begin{cases} 1 & x \in N \\ -1 & x \in N^c \end{cases}$$

Then f is non-measurable but $|f|$ is the constant function 1 which is measurable. $\square$

Proposition 2.4.8. *If $\{f_i\}$ are a sequence of measurable functions then $\sup_i f_i(x)$, $\inf_i f_i(x)$, $\limsup_{i\to\infty} f_i(x)$ and $\liminf_{i\to\infty} f_i(x)$ are all measurable. Consequently, $f(x) := \lim f_i(x)$, if exists, is measurable.*

Proof. If $f(x) := \sup_i f_i(x)$ then $\{f > a\} = \cup_i \{f_i > a\}$ and is measurable. If $f(x) := \inf_i f_i(x)$ then $f(x) = -\sup_i(-f_i(x))$. Also, $\limsup_{i\to\infty} f_i(x) = \inf_j(\sup_{i\geq j} f_i)$ and $\liminf_{i\to\infty} f_i(x) = \sup_j(\inf_{i\geq j} f_i)$. □

The space of measurable functions $M(\mathbb{R}^n)$ is closed under point-wise convergence. If $C(\mathbb{R}^n)$ denotes the space of all continuous functions on $\mathbb{R}^n$, then we have already seen that $C(\mathbb{R}^n) \subset M(\mathbb{R}^n)$. We know from classical analysis that $C(\mathbb{R}^n)$ is not closed under point-wise convergence and $M(\mathbb{R}^n)$ can be thought of as the "completion" of $C(\mathbb{R}^n)$ under point-wise convergence.

Recall that in Riemann integration, we approximated the graph of a given function by polygons, equivalently, we were approximating the given function by step functions. We shall now introduce a general class of functions which includes the step functions which corresponds to Lebesgue intergation.

Definition 2.4.9. *A finite linear combination of characteristic functions is called a* simple function*, i.e., a function $\phi : E \subset \mathbb{R}^n \to \mathbb{R}$ is said to be a* simple function *if it is of the form*

$$\phi(x) = \sum_{i=1}^{k} a_i \chi_{E_i}$$

for measurable subsets $E_i \subset \mathbb{R}^n$ with $\mu(E_i) < +\infty$ and $a_i \in \mathbb{R}$, for all i. A simple function ϕ is said to be a step function *if $E_i = R_i$ are the (bounded) cells in $\mathbb{R}^n$.*

By definition, a simple function is measurable and finite, hence in $M(\mathbb{R}^n)$. The class of a simple functions forms a vector subspace of $M(\mathbb{R}^n)$ as seen from the exercise below.

Exercise 39. If ϕ and ψ are simple functions on $\mathbb{R}^n$ then $\phi + \psi$ and $\phi\psi$ are simple too. Also, if ϕ is simple, $\lambda\phi$ is simple for all $\lambda \in \mathbb{R}$.

Note that the representation of the simple function ϕ, by our definition, is not unique.

Definition 2.4.10. *A non-zero simple function ϕ is said to have the* canonical representation *if*

$$\phi(x) = \sum_{i=1}^{k} a_i \chi_{E_i}$$

for disjoint measurable subsets $E_i \subset \mathbb{R}^n$ with $\mu(E_i) < +\infty$ and $a_i \neq 0$, for all i, and $a_i \neq a_j$ for $i \neq j$.

Exercise 40. Every non-zero simple function can be decomposed in to its unique canonical representation.

Proof. Let ϕ be a simple function. Then ϕ can take only finitely many distinct values. Let $\{b_1, \dots, b_k\}$ be the distinct non-zero values attained by ϕ. Define $E_i := \{x \in \mathbb{R}^n \mid \phi(x) = b_i\}$. By definition, E_i's are disjoint and we have the canonical representation of ϕ. $\square$

Exercise 41. If ϕ is simple with canonical representation $\phi = \sum_i a_i \chi_{E_i}$ then $|\phi|$ is simple and $|\phi| = \sum_i |a_i| \chi_{E_i}$.

Theorem 2.4.11. *For any finite a.e. measurable function f on $\mathbb{R}^n$ such that $f \geq 0$, there exists a sequence of simple functions $\{\phi_k\}_1^\infty$ such that*

(i) $\phi_k \geq 0$, *for each k, (non-negative)*

(ii) $\phi_k(x) \leq \phi_{k+1}(x)$ *(increasing sequence) and*

(iii) $\lim_{k\to\infty} \phi_k(x) = f(x)$ *for all x (point-wise convergence).*

Proof. Note that the domain of the given function may be of infinite measure. We begin by assuming that f is bounded, $|f(x)| \leq M$, and the support of f is contained in a cell R_M of equal side length M (a cube) centred at origin[9].

We now partition the range of f, $[0, M]$ in the following way: at every stage $k \geq 1$, we partition the range $[0, M]$ with intervals of length $1/2^k$ and correspondingly define the set,

$$E_{i,k} = \left\{ x \in R_M \mid \frac{i}{2^k} \leq f(x) \leq \frac{i+1}{2^k} \right\} \qquad \text{for all integers } 0 \leq i < M2^k.$$

Thus, for every $k \geq 1$, we have a disjoint partition of the domain of f, R_M, in to $\{E_{i,k}\}_i$. $E_{i,k}$ are all measurable due to the measurability of f. Hence, for each k, we define the simple function

$$\phi_k(x) = \sum_{i \in I_k} \frac{i}{2^k} \chi_{E_{i,k}}(x),$$

where I_k is the set of all integers in $[0, M2^k)$. By definition ϕ_k's are non-negative and $\phi_k(x) \leq f(x)$ for all $x \in R_M$ and for all k. In particular, $|\phi_k| \leq M$ for all k.

[9]The reason being a simple function is supported on finite measure set

We shall now show that ϕ_k's are an increasing sequence. Fix k and let $x \in R_M$. Then $x \in E_{j,k}$, for some $j \in I_k$, and $\phi_k(x) = j/2^k$. Similarly, there is a $j' \in I_{k+1}$ and $\phi_{k+1}(x) = j'/2^{k+1}$. The way we chose our partition, we know that $j' = 2j$ or $2j+1$. Therefore, $\phi_k(x) \leq \phi_{k+1}(x)$. Also, by definition of ϕ_k, $\phi_k(x) \leq f(x)$ for all $x \in R_M$.

It now remains to show the convergence. Now, for each $x \in R_M$,

$$|f(x) - \phi_k(x)| \leq \left| \frac{j+1}{2^k} - \frac{j}{2^k} \right| = \frac{1}{2^k}.$$

Thus, we have the convergence.

For any general non-negative measurable function, we construct a sequence of bounded functions f_k, $|f_k(x)| \leq k$, supported on a set of finite measure. Consider a cell R_k of equal side length k (a cube) centred at the origin. We define the truncation of f at k level on R_k, for each k, as follows:

$$f_k(x) = \begin{cases} f(x) & \text{if } x \in R_k \text{ and } f(x) \leq k \\ k & \text{if } x \in R_k \text{ and } f(x) > k \\ 0 & \text{ elsewhere.} \end{cases}$$

By construction, $f_k(x) \leq f_{k+1}(x)$, for all $x \in \mathbb{R}^n$. Also, $f_k(x) \overset{k\to\infty}{\longrightarrow} f(x)$ converges point-wise, for all $x \in \mathbb{R}^n$. To see this fact, fix $x \in \mathbb{R}^n$ and let ℓ be the such that $x \notin R_k$ for all $k < \ell$. Thus, $f_k(x) = 0$ for all $k = 1, 2, \ldots, \ell - 1$. Let $m = f(x)$. If $m \leq \ell$ then $f_k(x) = f(x)$ for all $k \geq \ell$ and hence the sequence converges point-wise. If $m > \ell$, choose the first integer i such that $\ell + i > m > \ell$ and we have $f_k(x) = f(x)$ for all $k \geq \ell + i$, and converges to $f(x)$.

Note that each f_k is measurable due to the measurability of f. Since the range of f_k is $[0, k]$, it is enough to check the measurablility of f_k for all $\alpha \in [0, k]$. The extreme cases, $\{f_k \leq 0\} = R_k^c$, $\{f_k < 0\} = \emptyset$ and $\{f_k \leq k\} = \mathbb{R}^n$ are measurable. For any $\alpha \in (0, k)$, $\{f_k \leq \alpha\} = R_k^c \cup \{f \leq \alpha\}$ is measurable.

For each k, f_k is non-negative bounded measurable function supported on a set of finite measure. Thus, for each k, we have a sequence of simple functions $\psi_{k\ell}$ satisfying the required properties and $\psi_{k\ell} \to f_k$, as $\ell \to \infty$. We pick the diagonal sequence $\phi_k = \psi_{kk}$. Note that ϕ_k is increasing sequence because $\psi_{k\ell}(x) \leq \psi_{(k+1)\ell}$. Also,

$$|\phi_k(x) - f(x)| \leq |\psi_{kk}(x) - f_k(x)| + |f_k(x) - f(x)| \leq \frac{1}{2^k} + |f_k(x) - f(x)|.$$

Thus, we have the point-wise convergence for all x. $\square$

Exercise 42. Show that in the result proved above if, in addition, f is bounded ($f(x) \leq M$) then the convergence is uniform.

Exercise 43. Let f be a non-negative measurable function on $\mathbb{R}^n$. Show that there exists a sequence of measurable subsets $\{E_k\}$ of $\mathbb{R}^n$ such that

$$f = \sum_{k=1}^{\infty} \frac{1}{k} \chi_{E_k}.$$

Proof. Let $x \in \mathbb{R}^n$ be such that $x \in \{f = 0\}$. Then we need to define E_k such that

$$\sum_{k=1}^{\infty} \frac{1}{k} \chi_{E_k}(x) = 0.$$

Equivalently, we need to define E_k such that $x \notin E_k$ for all k. Thus, $\{f = 0\} \cap E_k = \emptyset$ for every k. This suggests that on E_k, for every k, f is strictly positive. Also, if $x \in \mathbb{R}^n$ is such that $x \in E_k$ for all k, then $f(x) = +\infty$. Let $E_1 = \{x \in \mathbb{R}^n \mid f(x) \geq 1\}$ and, for $k = 2, 3, \ldots$, we define

$$E_k = \left\{ x \in \mathbb{R}^n \mid f(x) \geq \frac{1}{k} + \sum_{i=1}^{k-1} \frac{1}{i} \chi_{E_i}(x) \right\}.$$

By construction,

$$f(x) \geq \sum_{k=1}^{\infty} \frac{1}{k} \chi_{E_k}.$$

This is because at every stage k,

$$f(x) \geq \sum_{i=1}^{k} \frac{1}{i} \chi_{E_i}(x)$$

Clearly, the equality is true for $\{f = 0\}$ and $\{f = +\infty\}$. Now, fix $x \in \mathbb{R}^n$ such that $0 < f(x) < +\infty$, then $x \notin E_{k_m}$ for a subsequence k_m of k. Consequently,

$$f(x) < \frac{1}{k_m} + \sum_{i=1}^{k_m - 1} \frac{1}{i} \chi_{E_i}(x).$$

Letting $k_m \to \infty$, we have $f(x) < \sum_{k=1}^{\infty} \frac{1}{k} \chi_{E_k}(x)$. Hence, the equality holds. $\square$

In the next result, we relax the non-negativity requirement on f.

Theorem 2.4.12. *For any finite a.e. measurable function f on $\mathbb{R}^n$ there exists a sequence of simple functions $\{\Phi_k\}_1^\infty$ such that*

(i) $|\Phi_k(x)| \leq |\Phi_{k+1}(x)|$ *and*

(ii) $\lim_{k\to\infty} \Phi_k(x) = f(x)$ *for all x (point-wise convergence).*

In particular, $|\Phi_k(x)| \leq |f(x)|$ for all x and k.

Proof. Any function f can be decomposed in to non-negative functions as follows: $f = f^+ - f^-$. Corresponding to each we have a sequence of ϕ_k and ψ_k satisfying properties of previous theorem and converges point-wise to f^+ and f^-, respectively. By setting, $\Phi_k := \phi_k - \psi_k$ we immediately see that $\Phi_k(x) \to f(x)$ for all x. Let $E_1 := \{f < 0\}$, $E_2 := \{f > 0\}$ and $E_3 := \{f = 0\}$. Since both f^+ and f^- vanishes on E_3, ϕ_k, ψ_k vanishes on E_3. Thus, $\Phi_k = 0$ on E_3. Similarly, on E_1, $\Phi_k = -\psi_k \leq 0$ and on E_2, $\Phi_k = \phi_k$. Therefore, $|\Phi_k| = \phi_k + \psi_k$ and hence is an increasing sequence. □

In the above two theorems, we may allow extended real valued measurable function f, provided we allow the point-wise limit to take $\pm\infty$. The results above shows the density of simple functions in the space of finite valued functions in $M(\mathbb{R}^n)$ under the topology of point-wise convergence.

2.5 Littlewood's Three Principles

We have, thus far, developed the notion of measurable sets of $\mathbb{R}^n$ and measurable functions on $\mathbb{R}^n$. J. E. Littlewood[10] simplified the connection of theory of measures with classical real analysis in the following three observations:

(i) Every measurable set is "nearly" a finite union of intervals.

(ii) Every measurable function is "nearly" continuous.

(iii) Every convergent sequence of measurable functions is "nearly" uniformly convergent.

Littlewood had no contribution in the proof of these principles. He summarised the connections of measure theory notions with classical analysis.

[10] In his book on complex analysis titled *Lectures on the Theory of Functions*

2.5.1 First Principle

Theorem 2.5.1 (First Principle)**.** *If E is a measurable subset of $\mathbb{R}^n$ such that $\mu(E) < +\infty$ then, for every $\varepsilon > 0$, there exists a finite union of closed cells, say Γ, such that $\mu(E \triangle \Gamma) \leq \varepsilon$.*[11]

Proof. For every $\varepsilon > 0$, there exists a closed cover of cells $\{R_i\}_1^\infty$ for E ($E \subset \cup_{i=1}^\infty R_i$) such that

$$\sum_{i=1}^{\infty} |R_i| \leq \mu(E) + \frac{\varepsilon}{2}.$$

Since $\mu(E) < +\infty$, the series converges. Thus, for the given $\varepsilon > 0$ there exists a $k \in \mathbb{N}$ such that

$$\left| \sum_{i=1}^{k} |R_i| - \sum_{i=1}^{\infty} |R_i| \right| = \sum_{i=k+1}^{\infty} |R_i| < \frac{\varepsilon}{2}.$$

Set $\Gamma = \cup_{i=1}^k R_i$. Now,

$$\begin{aligned}
\mu(E\triangle\Gamma) &= \mu(E \setminus \Gamma) + \mu(\Gamma \setminus E) \quad \text{(by additivity)} \\
&\leq \mu\left(\cup_{i=k+1}^\infty R_i\right) + \mu\left(\cup_{i=1}^\infty R_i \setminus E\right) \quad \text{(by monotonicity)} \\
&= \mu\left(\cup_{i=k+1}^\infty R_i\right) + \mu\left(\cup_{i=1}^\infty R_i\right) - \mu(E) \quad \text{(by additivity)} \\
&\leq \sum_{i=k+1}^{\infty} |R_i| + \sum_{i=1}^{\infty} |R_i| - \mu(E) \quad \text{(by sub-additivity)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

Note that in the one dimension case one can, in fact, find a finite union of open intervals satisfying above condition.

At the end of last section, we saw that the sequence of simple functions were dense in $M(\mathbb{R}^n)$ under point-wise convergence. Using, Littlewood's first principle, one can say that the space of step functions is dense in $M(\mathbb{R}^n)$ under a.e. point-wise convergence.

[11] $E\triangle\Gamma = (E \cup \Gamma) \setminus (E \cap \Gamma)$

Theorem 2.5.2. *For any finite a.e. measurable function f on $\mathbb{R}^n$ there exists a sequence of step functions $\{\phi_k\}$ that converge to $f(x)$ point-wise for a.e. $x \in \mathbb{R}^n$.*

Proof. It is enough to show the claim for any characteristic function $f = \chi_E$, where E is a measurable subset of finite measure. By Littlewood's first principle, for every integer $k > 0$, there exist $\cup_{i=1}^{\ell} R_i$, finite union of closed cells such that $\mu(E \triangle \cup_{i=1}^{\ell} R_i) \leq 1/2^k$. It will be necessary to consider cells that are disjoint to make the value of simple function coincide with f in the intersection. Thus, we extend sides of R_i and form new collection of almost disjoint cells Q_i such that $\cup_{i=1}^{\ell} R_i = \cup_{i=1}^{m} Q_i$. Further still, for each integer $k > 0$, we can pick disjoint cells $P_i^k \subset Q_i$ such that $F_k := \cup_{i=1}^{m} P_i^k \subset \cup_{i=1}^{\ell} R_i$ and $\mu((\cup_{i=1}^{\ell} R_i) \setminus F_k) \leq 1/2^k$. Define

$$\phi_k := \chi_{F_k} = \sum_{i=1}^{m} \chi_{P_i^k}$$

is a step function and $f(x) = \phi_k(x)$ for all $x \in E \cap F_k$ and

$$\mu(\{x \in \mathbb{R}^n \mid f(x) \neq \phi_k\}) = \mu(E \triangle F_k) \leq \frac{1}{2^{k-1}}.$$

Observe that $G := \cap_{k=1}^{\infty} \cup_{i=k}^{\infty} (E \triangle F_k) = \{f \not\to \phi_k\}$ the set of all points where pointwise convergence fails. By first Borel-Cantelli (theorem 2.3.13), $\mu(G) = 0$ because $\sum_k (E \triangle F_k)$ is finite. $\square$

The space of step functions on $\mathbb{R}^n$ is dense in $M(\mathbb{R}^n)$ endowed with the point-wise a.e. topology.

2.5.2 Third Principle

The time is now ripe to state and prove Littlewood's third principle, a consequence of which is that uniform convergence is valid on "large" subset for a point-wise convergence sequence.

Theorem 2.5.3. *Suppose $\{f_k\}$ is a sequence of measurable functions defined on a measurable set E with $\mu(E) < +\infty$ such that $f_k(x) \to f(x)$ (point-wise) a.e. on E. Then, for any given $\varepsilon, \delta > 0$, there is a measurable subset $F_\delta^\varepsilon \subset E$ such that $\mu(E \setminus F_\delta^\varepsilon) < \delta$ and an integer $N \in \mathbb{N}$ (independent of x) such that, for all $x \in F_\delta^\varepsilon$,*

$$|f_k(x) - f(x)| < \varepsilon \quad \forall k \geq N.$$

Proof. We assume without loss of generality that $f_k(x) \to f(x)$ point-wise, for all $x \in E$. Otherwise we shall restrict ourselves to the subset of E where it holds and its complement in E is of measure zero.

For each $\varepsilon > 0$ and $x \in E$, there exists a $k \in \mathbb{N}$ (possibly depending on x) such that

$$|f_j(x) - f(x)| < \varepsilon \quad \forall j \geq k.$$

Since we want to get the region of uniform convergence, we accumulate all $x \in E$ for which the same k holds for a fixed ε. For the fixed $\varepsilon > 0$ and for each $k \in \mathbb{N}$, we define the set

$$E_k^\varepsilon := \{x \in E \mid |f_j(x) - f(x)| < \varepsilon \quad \forall j \geq k\}.$$

Note that not all of E_k^ε's are empty, otherwise it will contradict the point-wise convergence for all $x \in E$. Due to the measurability of f_k and f, E_k^ε are measurable. Also, note that by definition $E_k^\varepsilon \subset E_{k+1}^\varepsilon$ and $\cup_{k=1}^\infty E_k^\varepsilon = E$ (Exercise!). Thus, by continuity from below, for each $\delta > 0$, there is a $k_\delta \in \mathbb{N}$ such that

$$\mu(E) - \mu(E_k^\varepsilon) = \mu(E \setminus E_k^\varepsilon) < \delta \quad \forall k \geq k_\delta.$$

If $E_{k_\delta}^\varepsilon \neq \emptyset$, set $F_\delta^\varepsilon := E_{k_\delta}^\varepsilon$ and $N := k_\delta$ else set F_δ^ε to be the first non-empty set E_m^ε for $m \geq k_\delta$ and $N := m$. We have, in particular, $\mu(E \setminus F_\delta^\varepsilon) < \delta$ and for all $x \in F_\delta^\varepsilon$,

$$|f_j(x) - f(x)| < \varepsilon \quad \forall j \geq N.$$

□

Corollary 2.5.4. *Suppose $\{f_k\}$ is a sequence of measurable functions defined on a measurable set E with $\mu(E) < +\infty$ such that $f_k(x) \to f(x)$ (point-wise) a.e. on E. Then, for any given $\varepsilon, \delta > 0$, there is a closed subset $\Gamma_\delta^\varepsilon \subset E$ such that $\mu(E \setminus \Gamma_\delta^\varepsilon) < \delta$ and an integer $K \in \mathbb{N}$ (independent of x) such that, for all $x \in \Gamma_\delta^\varepsilon$,*

$$|f_k(x) - f(x)| < \varepsilon \quad \forall k \geq K.$$

Proof. Using above theorem obtain F_δ^ε such that $\mu(E \setminus F_\delta^\varepsilon) < \delta/2$. By inner regularity, pick a closed set $\Gamma_\delta^\varepsilon \subset F_\delta^\varepsilon$ such that $\mu(F_\delta^\varepsilon \setminus \Gamma_\delta^\varepsilon) < \delta/2$. □

Note that in the Theorem and Corollary above, the choice of the set F_δ^ε or $\Gamma_\delta^\varepsilon$ may depend on ε. We can, in fact, have a stronger result that one can choose the set independent of ε.

Corollary 2.5.5 (Egorov)**.** *Suppose $\{f_k\}$ is a sequence of measurable functions defined on a measurable set E with $\mu(E) < +\infty$ such that $f_k(x) \to f(x)$ (point-wise) a.e. on E. Then, for any given $\delta > 0$, there is a measurable subset $F_\delta \subset E$ such that $\mu(E \setminus F_\delta) < \delta$ and $f_k \to f$ uniformly on F_δ.*

Proof. From the theorem proved above, for a given $\delta > 0$ and $k \in \mathbb{N}$, there is a measurable subset $F_k \subset E$ such that $\mu(E \setminus F_k) < \delta/2^k$ and for all $x \in F_k$, there is a $N_k \in \mathbb{N}$

$$|f_j(x) - f(x)| < 1/k \quad \forall j > N_k.$$

Set $F_\delta := \cap_{k=1}^{\infty} F_k$. Thus,

$$\mu(E \setminus F_\delta) = \mu(\cup_{k=1}^{\infty}(E \setminus F_k)) \leq \sum_{k=1}^{\infty} \mu(E \setminus F_k) < \delta.$$

Now, it is easy to check that $f_k \to f$ uniformly in F_δ. □

Example 2.15. The finite measure hypothesis on E is necessary in above theorems. Let $f_k = \chi_{[k,k+1)}$ then $f_k(x) \to 0$ point-wise. Choose $\delta = 1$ and let F be any subset of $\mathbb{R}$ such that $\mu(F) < 1$. Because $\mu(F) < 1$, $F^c \cap [k, k+1) \neq \emptyset$, for all k. We claim that f_k cannot converge uniformly to 0 on F^c. For every $k \in \mathbb{N}$, there is a $x_k \in F^c \cap [k, k+1)$ such that $|f_k(x_k) - f(x_k)| = |f_k(x_k)| = 1$.

The Egorov's theorem motivates the following notion of convergence in $M(\mathbb{R}^n)$.

Definition 2.5.6. *Let $\{f_k\}_1^\infty$ and f be finite a.e. measurable functions on a measurable set $E \subseteq \mathbb{R}^n$. We say f_k* converges almost uniformly[12] *to f on E, if for every $\delta > 0$, there exists measurable subset $F_\delta \subset E$ such that $\mu(F_\delta) < \delta$ and $f_k \to f$ uniformly on $E \setminus F_\delta$.*

Exercise 44. Show that almost uniform convergence implies point-wise a.e. convergence.

Proof. Let $f_k \to f$ almost uniformly converge. Then, by definition, for each $k \in \mathbb{N}$, there exists a measurable set $F_k \subset E$ such that $\mu(F_k) < 1/k$ and $f_k \to f$ uniformly on $E \setminus F_k$. Let $F = \cap_{k=1}^{\infty} F_k$. Thus, $\mu(F) \leq \mu(F_k) < 1/k$ for all k and hence $\mu(F) = 0$. For any $x \in E \setminus F$, $x \in \cup_{k=1}(E \setminus F_k)$ and hence

[12]Note that this notion of convergence is much weaker than demanding uniform convergence except on zero measure sets.

$x \in E \setminus F_k$ for some k. Therefore, by the uniform convergence of $f_n \to f$, we have $f_n(x) \to f(x)$ point-wise. Thus, $f_n(x) \to f(x)$ for all $x \in E \setminus F$ and $\mu(F) = 0$, showing the point-wise a.e. convergence. □

The converse is not true. Example 2.15 gives an example of a point-wise a.e. converging sequence which do not converge almost uniformly. However, the converse is true for a finite measure set E. The Egorov's theorem is precisely the converse statement for finite measure set. Thus, on finite measure set we have the following statement:

Exercise 45. For finite measure set $\mu(E) < +\infty$, a sequence of functions on E converges point-wise a.e. iff it converges almost uniformly.

We shall end this section by giving a weaker notion of convergence on $M(\mathbb{R}^n)$.

Definition 2.5.7. *Let $\{f_k\}_1^\infty$ and f be finite a.e. measurable functions on a measurable set $E \subseteq \mathbb{R}^n$. We say f_k* converges in measure *to f on E, denoted as $f_k \xrightarrow{\mu} f$, if for every $\varepsilon > 0$,*

$$\lim_{k\to\infty} \mu(E_k^\varepsilon) = 0.$$

where

$$E_k^\varepsilon := \{x \in E \mid |f_k(x) - f(x)| > \varepsilon\}.$$

Exercise 46. Almost uniform convergence implies convergence in measure.

Proof. Since f_k converges almost uniformly to f. For every $\varepsilon > 0$ and $k \in \mathbb{N}$, there exists a set F_k (independent of ε) such that $\mu(F_k) < 1/k$ and there exists $K \in \mathbb{N}$, for all $x \in E \setminus F_k$, such that

$$|f_j(x) - f(x)| \leq \varepsilon \quad \forall j > K.$$

Therefore $E_j^\varepsilon \subset F_k$ for infinitely many $j > K$. Thus, for infinitely many j, $\mu(E_j^\varepsilon) < 1/k$. In particular, choose $j_k > k$ and $\mu(E_{j_k}^\varepsilon) < 1/k$. □

Example 2.16. Give an example to show that convergence in measure do not imply almost uniform convergence.

However, the converse is true upto a subsequence.

Theorem 2.5.8. *Let $\{f_k\}_1^\infty$ and f be finite a.e. measurable functions on a measurable set $E \subseteq \mathbb{R}^n$ (not necessarily finite). If $f_k \xrightarrow{\mu} f$ then there is a subsequence $\{f_{k_l}\}_{l=1}^\infty$ such that f_{k_l} converges almost uniformly to f.*

2.5.3 Second Principle

Recall $M(\mathbb{R}^n)$ is decomposed in to equivalence classes under equality a.e. So, it would be a nice situation if for every measurable function there is a continuous function in its equivalence class. In other words, we wish to have for every measurable function f a continuous function g such that $f = g$ a.e. Unfortunately, this is not true.

Exercise 47. Given an example of a measurable function f for which there is no continuous function g such that $f = g$ a.e.

Littlewood's second principle is "approximating" a measurable function on finite measure by a continuous function.

Exercise 48. Let χ_R be a step function on $\mathbb{R}^n$, where R is a cell with $|R| < +\infty$. Then, for $\varepsilon > 0$, there exists a subset $E_\varepsilon \subset \mathbb{R}^n$ such that χ_R restricted to $\mathbb{R}^n \setminus E_\varepsilon$ is continuous and $\mu(E_\varepsilon) < \varepsilon$.

Theorem 2.5.9 (Luzin)**.** *Let f be measurable finite a.e. on a measurable set E such that $\mu(E) < +\infty$. Then, for $\varepsilon > 0$, there exists a closed set $\Gamma_\varepsilon \subset E$ such that $\mu(E \setminus \Gamma_\varepsilon) < \varepsilon$ and $f\,|_{\Gamma_\varepsilon}$ is a continuous function*[13].

Proof. We know from Theorem 2.5.2 that step functions are dense in $M(\mathbb{R}^n)$ under point-wise a.e. convergence. Let $\{\phi_k\}_{k=1}^{\infty}$ be a sequence of step functions that converge to f point-wise a.e. Fix $\varepsilon > 0$. For each $k \in \mathbb{N}$, there exists a measurable subset $E_k \subset E$ such that $\mu(E_k) < (\varepsilon/3)(1/2^k)$ and ϕ_k restricted to $E \setminus E_k$ is continuous. By Egorov's theorem, there is a measurable set $F_\varepsilon \subset E$ such that $\mu(E \setminus F_\varepsilon) < \varepsilon/3$ and $\phi_k \to f$ uniformly in F_ε. Note that ϕ_k restricted to $G_\varepsilon := F_\varepsilon \setminus \cup_{k=1}^{\infty} E_k$ is continuous. Therefore, its uniform limit f restricted to G_ε is continuous. Also,

$$\mu(E \setminus G_\varepsilon) = \mu(\cup_{k=1}^{\infty} E_k) + \mu(E \setminus F_\varepsilon) < (2\varepsilon)/3.$$

By inner regularity, pick a closed set $\Gamma_\varepsilon \subset G_\varepsilon$ such that $\mu(G_\varepsilon \setminus \Gamma_\varepsilon) < \varepsilon/3$. Now, obviously, $f\,|_{\Gamma_\varepsilon}$ is continuous and

$$\mu(E \setminus \Gamma_\varepsilon) = \mu(E \setminus G_\varepsilon) + \mu(G_\varepsilon \setminus \Gamma_\varepsilon) < \varepsilon.$$

□

[13] f restricted to Γ_ε is continuous but f as a function on E may not be continuous on points of Γ_ε

Corollary 2.5.10. *Let f be measurable finite a.e. on a measurable set E such that $\mu(E) < +\infty$. Then, for $\varepsilon > 0$, there exists a continuous function g on E such that*

$$\mu(\{x \in E \mid f(x) \neq g(x)\} < \varepsilon.$$

Proof. Use Urysohn lemma or Tietze extension theorem to find a continuous function g on E which coincides with f on Γ_ε. Consider the set

$$\{x \in E \mid f(x) \neq g(x)\} = E \setminus \Gamma_\varepsilon.$$

The measure of the above set is less than ε. $\square$

Exercise 49. For any finite a.e. measurable function f on $\mathbb{R}^n$ there exists a sequence of *continuous* functions $\{f_k\}$ that converge to $f(x)$ point-wise for a.e. $x \in \mathbb{R}^n$.

2.6 Jordan Content or Measure

We end this chapter with few remarks on the notion of "Jordan content of a set", developed by Giuseppe Peano and Camille Jordan. This notion is related to Riemann integration in the same way as Lebesgue measure is related to Lebesgue integration. The Jordan content is the finite version of the Lebesgue measure.

Definition 2.6.1. *Let E be bounded subset of $\mathbb{R}^n$. We say that a finite family of cells $\{R_i\}_{i\in I}$ is a* finite covering *of E iff $E \subseteq \cup_{i\in I} R_i$, where I is a finite index set.*

Let $B \subset 2^{\mathbb{R}^n}$ be the class of all bounded subsets of $\mathbb{R}^n$. The reason for restricting ourselves to B is because every element of B will admit a finite covering.

Definition 2.6.2. *For a subset $E \in B$, we define its* Jordan outer content *$J^\star(E)$ as,*

$$J^\star(E) := \inf_{E \subseteq \cup_{i\in I} R_i} \sum_{i\in I} |R_i|,$$

the infimum being taken over all possible finite *coverings of E.*

The term "measure" is usually reserved for a countably additive set function, hence we use the term "content". Otherwise Jordan content could be viewed as finitely additive measure. Some texts refer to it as Jordan measure or Jordan-Peano measure.

Lemma 2.6.3. *The Jordan outer content $J^\star$ has the following properties:*

(a) For every subset $E \in B$, $0 \leq J^\star(E) < +\infty$.

*(b) (**Translation Invariance**) For every $E \in B$, $J^\star(E+x) = J^\star(E)$ for all $x \in \mathbb{R}^n$.*

*(c) (**Monotone**) If $E \subset F$, then $J^\star(E) \leq J^\star(F)$.*

*(d) (**Finite Sub-additivity**) For a finite index set I,*

$$J^\star\left(\cup_{i\in I} E_i\right) \leq \sum_{i\in I} J^\star(E_i).$$

Proof. The proofs are similar to those of outer measure case. □

Exercise 50. Show that $J^\star(R) = |R|$ for any cell $R \subset \mathbb{R}^n$. Consequently, $\mu^\star(R) = J^\star(R)$ for every cell R.

Exercise 51. If $E \subset \mathbb{R}^2$ denotes the region below the graph of a bounded function $f : [a, b] \to \mathbb{R}$. Show that the Jordan content of E is same as the Riemann upper sum of f.

Example 2.17. Jordan outer content of the empty set is zero, $\mu^\star(\emptyset) = 0$. Every cell is a cover for the empty set. Thus, infimum over the volume of all cells is zero.

Example 2.18. The Jordan outer content for a singleton set $\{x\}$ in $\mathbb{R}^n$ is zero. The same argument as for empty set holds except that now the infimum is taken over all cells containing x. Thus, for each $\varepsilon > 0$, one can find a cell R_ε such that $x \in R_\varepsilon$ and $|R_\varepsilon| \leq \varepsilon$. Therefore, $\mu^\star(\{x\}) \leq \varepsilon$ for all $\varepsilon > 0$ and hence $\mu^\star(\{x\}) = 0$.

Example 2.19. The Jordan outer content of a finite subset E of $\mathbb{R}^n$ is zero. A finite set $E = \cup_{x\in E}\{x\}$, where the union is finite. Thus, by finite sub-additivity, $\mu^\star(E) \leq 0$ and hence $\mu^\star(E) = 0$.

This is precisely where the difference lies between Lebesgue measure and Jordan content. Recall the $\mathbb{Q}$ had Lebesgue outer measure zero. But the Jordan outer content of $\mathbb{Q}$ contained in a bounded cell is positive. For instance, consider $E := \mathbb{Q} \cap [0, 1]$. We shall show that $J^\star(E) = 1$. Note that there is no finite cover of E which is properly contained in $[0, 1]$, due to the density of $\mathbb{Q}$ in $[0, 1]$. Thus, any finite cover of E also contains $[0, 1]$. In fact, $[0, 1]$ is itself a finite cover of E. Infimum over all the finite cover is bounded below by 1, due to monotonicity. Thus, $J^\star(E) = 1$.

Exercise 52. Show that $J^\star(E) = J^\star(\overline{E})$, where $\overline{E}$ denotes the closure of E.

Proof. Firstly, the result is true when $E = R$ is a cell, since $J^\star(\overline{R}) = \mu^\star(\overline{R}) = \mu^\star(R) = J^\star(R)$. By monotonicity of $J^\star$, $J^\star(E) \leq J^\star(\overline{E})$. For converse argument, let $\{R_i\}$ be any finite cover of E, i.e., $E \subset \cup_i R_i$ then $\overline{E} \subset \overline{\cup_i R_i}$. In general, $\cup_i \overline{R_i} \subseteq \overline{\cup_i R_i}$. However, since the union is finite we have equality,

$$\cup_i \overline{R_i} = \overline{\cup_i R_i}.$$

Therefore, $\{\overline{R}_i\}$ is a finite cover of $\overline{E}$. Thus,

$$J^\star(\overline{E}) \leq \sum_i |\overline{R}_i| = \sum_i |R_i|.$$

Taking infimum over all finite covers of E, we get $J^\star(\overline{E}) \leq J^\star(E)$. □

Note that in the proof above the finite cover played a crucial role. A similar result is not true Lebesgue outer measure. For instance, $\mu^\star(\mathbb{Q} \cap [0, 1]) = 0$ and its closure is $[0, 1]$ whose outer measure in one. More generally, for every $k \geq 0$, we have a set $E \subset \mathbb{R}^n$ such that, $\mu^\star(E) = 0$ but $\mu^\star(\partial E) = k$, where ∂E is the boundary of E. For instance, consider $E := \mathbb{Q}^n \cap [0, k^{1/n}]^n$. Then, $\partial E = [0, k^{1/n}]^n$ and $\mu^\star(\partial E) = k$.

Exercise 53. Show that the Jordan content of a set is same as the outer measure of its closure, i.e., $\mu^\star(\overline{E}) = J^\star(E)$.

Proof. Note that it is enough to show that $\mu^\star(\overline{E}) = J^\star(\overline{E})$. Firstly, it follows from defintion that $\mu^\star(\overline{E}) \leq J^\star(\overline{E})$. For the reverse inequality, we consider $\{R_i\}$ to be an countable cover of $\overline{E}$. Since $\overline{E}$ is closed and bounded, hence compact, there is a finite sub-cover $\{Q_j\}$ of $\overline{E}$. Thus,

$$J^\star(\overline{E}) \leq \sum_j |Q_j| \leq \sum_i |R_i|.$$

Taking infimum over all countable covers $\{R_i\}$, we get

$$J^\star(\overline{E}) \leq \mu^\star(\overline{E}).$$

Hence the equality holds. $\square$

To classify the *Jordan measurable* subsets, we need to identify the class of subsets of B for which finite additivity holds. Note that $\mathbb{Q} \cap [0,1]$ cannot be Jordan measurable. Since $J^\star(\mathbb{Q} \cap [0,1]) = 1$. A similar argument also shows that $J^\star(\mathbb{Q}^c \cap [0,1] = 1$. If finite additivity were true then $J^\star(\mathbb{Q} \cap [0,1]) + J^\star(\mathbb{Q}^c \cap [0,1]) = 2 \neq 1 = J^\star([0,1])$.

Theorem 2.6.4. *A bounded set $E \subset \mathbb{R}^n$ is Jordan measurable iff χ_E is Riemann integrable.*

Now do you see why the characteristic function on $\mathbb{Q}$ was not Riemann integrable? Precisely because $\mathbb{Q}$ was not Jordan measurable. Do you also see how Lebesgue measure fixes this inadequacy?

Chapter 3

Lebesgue Integration

In this chapter, we shall define the integral of a function on $\mathbb{R}^n$, in a progressive way, with increasing order of complexity. Before we do so, we shall state some facts about Riemann integrability in measure theoretic language.

Theorem 3.0.1. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then $f \in \mathcal{R}([a, b])$ iff f is continuous a.e. on $[a, b]$*

Basically, the result says that a function is Riemann integrable iff its set of discontinuities are of length (measure) zero.

3.1 Simple Functions

Recall from the discussion on simple functions in previous chapter that the representation of simple functions is not unique. Therefore, we defined the canonical representation of a simple function which is unique. We use this canonical representation to define the integral of a simple function.

Definition 3.1.1. *Let ϕ be a non-zero simple function on $\mathbb{R}^n$ having the canonical form*

$$\phi(x) = \sum_{i=1}^{k} a_i \chi_{E_i}$$

with disjoint measurable subsets $E_i \subset \mathbb{R}^n$ with $\mu(E_i) < +\infty$ and $a_i \neq 0$, for all i, and $a_i \neq a_j$ for $i \neq j$. We define the Lebesgue integral *of a simple* *function on $\mathbb{R}^n$, denoted as*

$$\int_{\mathbb{R}^n} \phi(x)\, d\mu := \sum_{i=1}^{k} a_i \mu(E_i).$$

where μ is the Lebesgue measure on $\mathbb{R}^n$. Henceforth, we shall denote $\int \phi\, d\mu$ as $\int \phi\, dx$, for Lebesgue measure. Also, we define the integral of ϕ on $E \subset \mathbb{R}^n$ as,

$$\int_E \phi(x)\, dx := \int_{\mathbb{R}^n} \phi(x)\chi_E(x)\, dx.$$

Remark 3.1.2. Consider the zero function as the characteristic function $\chi_\emptyset$. Then integral of the zero function is defined as $\mu(\emptyset) = 0$.

Note that the integral of a simple function is always finite. Though, we chose to define integral using the canonical representation, it turns out that integral of a simple function is independent of its representation.

Proposition 3.1.3. *For any representation of the simple function $\phi = \sum_{i=1}^k a_i\chi_{E_i}$, we have*

$$\int_{\mathbb{R}^n} \phi(x)\, dx = \sum_{i=1}^{k} a_i \mu(E_i).$$

Proof. Let $\phi = \sum_{i=1}^k a_i\chi_{E_i}$ be a representation of ϕ such that E_i's are pairwise disjoint which is not the canonical form, i.e., a_i are not necessarily distinct and can be zero for some i. Let $\{b_j\}$ be the distinct non-zero elements of $\{a_1, \ldots, a_k\}$, where $1 \leq j \leq k$. For a fixed j, we define $F_j = \cup_{i\in I_j} E_i$, where $I_j := \{i \mid a_i = b_j\}$. Then F_j's are pairwise disjoint and $\mu(F_j) = \sum_{i\in I_j} \mu(E_i)$. Therefore, $\phi = \sum_j b_j\chi_{F_j}$ is a canonical form of ϕ. Thus,

$$\int_{\mathbb{R}^n} \phi(x)\, dx = \sum_j b_j\mu(F_j) = \sum_j b_j \sum_{i\in I_j} \mu(E_i) = \sum_{i=1}^{k} a_i\mu(E_i).$$

We now consider a representation of ϕ such that E_i are not necessarily disjoint. Let $\phi = \sum_{i=1}^k a_i\chi_{E_i}$ be any general representation of ϕ, $a_i \in \mathbb{R}$. Given any collection of subsets $\{E_i\}_1^k$ of $\mathbb{R}^n$, there exists a collection of disjoint subsets $\{F_j\}_1^m$, for $m \leq 2^k$, such that $\cup_j F_j = \mathbb{R}^n$, $\cup_{j=1}^{m-1} F_j = \cup_i E_i$ and, for each i, $E_i = \cup_{j\in I_i} F_j$ where $I_i := \{j \mid F_j \subset E_i\}$ (Exercise!). For each j, we define $b_j := \sum_{i \in I_j} a_i$ where $I_j := \{i \mid F_j \subset E_i\}$. Thus, $\phi = \sum_{j=1}^m b_j \chi_{F_j}$, where F_j are pairwise disjoint. Hence, from first part of the proof,

$$\begin{aligned}\int_{\mathbb{R}^n} \phi(x)\,dx &= \sum_{j=1}^m b_j \mu(F_j) = \sum_{j=1}^m \sum_{i \in I_j} a_i \mu(F_j) \\ &= \sum_{i=1}^k \sum_{j \in I_i} a_i \mu(F_j) = \sum_{i=1}^k a_i \mu(E_i).\end{aligned}$$

Hence the integral is independent of the choice of the representation. □

Exercise 54. Show the following properties of integral of simple functions:

(i) (Linearity) For any two simple functions ϕ, ψ and $\alpha, \beta \in \mathbb{R}$,

$$\int_{\mathbb{R}^n} (\alpha\phi + \beta\psi)\,dx = \alpha \int_{\mathbb{R}^n} \phi\,dx + \beta \int_{\mathbb{R}^n} \psi\,dx.$$

(ii) (Additivity) For any two disjoint subsets $E, F \subset \mathbb{R}^n$ with finite measure

$$\int_{E \cup F} \phi\,dx = \int_E \phi\,dx + \int_F \phi\,dx.$$

(iii) (Monotonicity) If $\phi \geq 0$ then $\int \phi \geq 0$. Consequently, if $\phi \leq \psi$, then

$$\int_{\mathbb{R}^n} \phi\,dx \leq \int_{\mathbb{R}^n} \psi\,dx.$$

(iv) (Triangle Inequality) We know for a simple function ϕ, $|\phi|$ is also simple. Thus,

$$\left| \int_{\mathbb{R}^n} \phi\,dx \right| \leq \int_{\mathbb{R}^n} |\phi|\,dx.$$

(v) If $\phi = \psi$ a.e. then $\int \phi = \int \psi$.

Example 3.1. An example of a Lebesgue integrable function which is not Riemann integral is the following: Consider the characteristic function $\chi_{\mathbb{Q}}$. We have already seen in Example 1.3 that this is not Riemann integrable. But

$$\int_{\mathbb{R}^n} \chi_{\mathbb{Q}}(x)\,dx = \mu(\mathbb{Q}) = 0.$$

However, $\chi_{\mathbb{Q}} = 0$ a.e. and zero function is Riemann integrable. Thus, for $\chi_{\mathbb{Q}}$ which is Lebesgue integrable function there is a Riemann integrable function in its equivalence class. Is this always true? Do we always have a Riemann integrable function in the equivalence class of a Lebesgue integrable function. The answer is a "no". Find an example!

Exercise 55. Show that the Riemann integral and Lebesgue integral coincide for step functions.

3.2 Bounded Function With Finite Measure Support

Now that we have defined the notion of integral for a simple function, we intend to extend this notion to other measurable functions. At this juncture, the natural thing is to recall the fact proved in Theorem 2.4.12, which establishes the existence of a sequence of simple functions ϕ_k converging point-wise to a given measurable finite a.e. function f. Thus, the natural way of defining the integral of the function f would be

$$\int_{\mathbb{R}^n} f(x)\,dx := \lim_{k\to\infty} \int_{\mathbb{R}^n} \phi_k(x)\,dx.$$

This definition may not be well-defined. For instance, the limit on the RHS may depend on the choice of the sequence of simple functions ϕ_k.

Example 3.2. Let $f \equiv 0$ be the zero function. By choosing $\phi_k = \chi_{(0,1/k)}$ which converges to f point-wise its integral is $1/k$ which also converging to zero. However, if we choose $\psi_k = k\chi_{(0,1/k)}$ which converges point-wise to f, but $\int \psi_k = 1$ for all k and hence converges to 1.

But zero function is trivially a simple function with Lebesgue integral zero. Note that the situation is very similar to what happens in Riemann's notion of integration. Therein we demand that the Riemann upper sum and Riemann lower sum coincide, for a function to be Riemann integrable. In Lebesgue's situation too, we have that the integral of different sequences of simple functions converging to a function f may not coincide. The following result singles out a case when the limits of integral of simple functions coincide for any choice.

Proposition 3.2.1. *Let f be a measurable function finite a.e. on a set E of finite measure and let $\{\phi_k\}$ be a sequence of simple functions supported on E*

and uniformly bounded by M such that $\phi_k(x) \to f(x)$ point-wise a.e. on E. Then $L := \lim_{k\to\infty} \int_E \phi_k\,dx$ is finite. Further, L is independent of the choice of $\{\phi_k\}$, i.e., if $f = 0$ a.e. then $L = 0$.

Proof. Since $\phi_k(x) \to f(x)$ point-wise a.e. on E and $\mu(E) < +\infty$, by Egorov's theorem, for a given $\delta > 0$, there exists a measurable subset $F_\delta \subset E$ such that $\mu(E \setminus F_\delta) < \delta/(4M)$ and $\phi_k \to f$ uniformly on F_δ. Set $I_k := \int_E \phi_k$. We shall show that $\{I_k\}$ is a Cauchy sequence in $\mathbb{R}$ and hence converges. Consider

$$
\begin{aligned}
|I_k - I_m| &\le \int_E |\phi_k(x) - \phi_m(x)| \quad \text{(triangle inequality)}\\
&= \int_{F_\delta} |\phi_k(x) - \phi_m(x)| + \int_{E\setminus F_\delta} |\phi_k(x) - \phi_m(x)|\\
&< \int_{F_\delta} |\phi_k(x) - \phi_m(x)| + \frac{\delta}{2}\\
&< \mu(F_\delta)\frac{\delta}{2\mu(E)} + \frac{\delta}{2} \quad \text{for all } k, m > K\\
&\le \delta \quad \text{for all } k, m > K \quad \text{(By monotonicity of } \mu).
\end{aligned}
$$

Thus, $\{I_k\}$ is Cauchy sequence and converges to some L. If $f = 0$, repeating the above argument on I_k

$$
\begin{aligned}
|I_k| &\le \int_E |\phi_k(x)| \quad \text{(triangle inequality)}\\
&= \int_{F_\delta} |\phi_k(x)| + \int_{E\setminus F_\delta} |\phi_k(x)|\\
&< \int_{F_\delta} |\phi_k(x)| + \frac{\delta}{4}\\
&< \mu(F_\delta)\frac{3\delta}{4\mu(E)} + \frac{\delta}{2} \quad \text{for all } k > K'\\
&\le \delta \quad \text{for all } k > K' \quad \text{(By monotonicity of } \mu),
\end{aligned}
$$

we get $L = 0$. □

We know from Theorem 2.4.12 the existence of a sequence of simple functions ϕ_k converging point-wise to a given measurable finite a.e. function f. If, in addition, we assume f is bounded and supported on a finite measure set E, then the ϕ_k satisfy the hypotheses of above Proposition. This motivates us to give the following definition.

Definition 3.2.2. *Let f be bounded measurable function supported on a set E of finite measure. The* integral *of f is defined as*

$$\int_E f(x)\,dx := \lim_{k\to\infty} \int_E \phi_k(x)\,dx,$$

where $\{\phi_k\}$ are uniformly bounded simple functions supported on the support of f and converging point-wise to f. Moreover, for any measurable subset $F \subset E$,

$$\int_F f(x)\,dx := \int_E f(x)\chi_F(x)\,dx.$$

Exercise 56. Show that all the properties of integral listed in Exercise 54 is also valid for an integral of a bounded measurable function with support on finite measure.

Exercise 57. A consequence of (v) property is that if $f = 0$ a.e. then $\int f = 0$. The converse is true for non-negative functions. Let f be a bounded measurable function supported on finite measure set. If $f \geq 0$ and $\int f = 0$ then $f = 0$ a.e.

The way we defined our integral of a function, the interchange of limit and integral under point-wise convergence comes out as a gift.

Theorem 3.2.3 (Bounded Convergence Theorem[1]). *Let f_k be a sequence of measurable functions supported on a finite measure set E such that $|f_k(x)| \leq M$ for all k and $x \in E$ and $f_k(x) \to f(x)$ point-wise a.e. on E. Then f is also bounded and supported on E a.e. and*

$$\lim_{k\to\infty} \int_E |f_k - f| = 0.$$

In particular,

$$\lim_{k\to\infty} \int_E f_k = \int_E f.$$

Proof. Since f is a point-wise a.e. limit of f_k, $|f(x)| \leq M$ a.e. on E and has support in E. By Egorov's theorem, for a given $\delta > 0$, there exists a measurable subset $F_\delta \subset E$ such that $\mu(E \setminus F_\delta) < \delta/(4M)$ and $f_k \to f$

[1] This statement is same as the one in Theorem 1.1.6 where we mentioned the proof is not elementary

uniformly on F_δ. Also, choose $K \in \mathbb{N}$, such that $|f_k(x) - f(x)| < \delta/2\mu(E)$ for all $k > K$. Consider

$$\begin{aligned}
\int_E |f_k(x) - f(x)| &\leq \int_{F_\delta} |f_k(x) - f(x)| + \int_{E\setminus F_\delta} |f_k(x) - f(x)| \\
&< \mu(F_\delta)\frac{\delta}{2\mu(E)} + \frac{\delta}{2} \quad \text{for all } k > K \\
&\leq \delta \quad \text{for all } k > K \quad \text{(By monotonicity of } \mu).
\end{aligned}$$

Therefore, $\lim_{k\to\infty} \int_E |f_k - f| = 0$ and, by triangle inequality,

$$\lim_{k\to\infty} \int_E f_k = \int_E f.$$

□

It is now time to address the problem of Riemann integration which does not allow us to interchange point-wise limit and integral. We first observe that Riemann integration is same as Lebesgue integration for Riemann integrable functions and thus, by BCT, we have the interchange of limit and integral for Riemann integrable functions, when the limit is also Riemann integrable.

Theorem 3.2.4. *If $f \in \mathcal{R}([a,b])$ then f is bounded measurable and*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f(x)\,dx,$$

the LHS is in the sense of Riemann and RHS in the sense of Lebesgue.

Proof. Since $f \in \mathcal{R}([a,b])$, f is bounded, $|f(x)| \leq M$ for some $M > 0$. Also, the support of f, being subset of $[a,b]$, is finite. We need to check that f is measurable. Since f is Riemann integrable there exists two sequences of step functions $\{\phi_k\}$ and $\{\psi_k\}$ such that

$$\phi_1 \leq \ldots \leq \phi_k \leq \ldots \leq f \leq \psi_k \leq \ldots \leq \psi_1$$

and

$$\lim_k \int_a^b \phi_k = \lim_k \int_a^b \psi_k = \lim_k f.$$

Also, $|\phi_k| \leq M$ and $|\psi_k| \leq M$ for all k. Since Riemann integral is same as Lebesgue integral for step functions,

$$\int_a^b \phi_k = \int_{[a,b]} \phi_k \text{ and } \quad \int_a^b \psi_k = \int_{[a,b]} \psi_k.$$

Let $\Phi(x) := \lim_k \phi_k(x)$ and $\Psi := \lim_k \psi_k(x)$. Thus, $\Phi \leq f \leq \Psi$. Being limit of simple functions Φ and Ψ are measurable and by BCT,

$$\int_{[a,b]} \Phi = \lim_k \int_a^b \phi = \lim_k \int_a^b \psi = \int_{[a,b]} \Psi.$$

Thus, $\int_{[a,b]}(\Psi - \Phi) = 0$. Moreover, since $\psi_k - \phi_k \geq 0$, we must have $\Psi - \Phi \geq 0$. Thus, by Exercise 57, $\Psi - \Phi = 0$ a.e. and hence $\Phi = f = \Psi$ a.e. Hence f is measurable. Thus,

$$\int_{[a,b]} f(x)\,dx = \lim_k \int_{[a,b]} \phi_k = \int_a^b f(x)\,dx.$$

□

The same statement is not true, in general, for improper Riemann integration (cf. Exercise 62).

3.3 Non-negative Functions

We have already noted that (cf. Example 3.2) for a general measurable function, defining its integral as the limit of the simple functions converging to it, may not be well-defined. However, we know from the proof of Theorem 2.4.11 that any non-negative function f has truncation f_k which are each bounded and supported on a set of finite measure, increasing and converge point-wise to f. There could be many other choices of the sequences which satisfy similar condition. This motivates a definition of integrability for non-negative functions.

Definition 3.3.1. *Let f be a non-negative ($f \geq 0$) measurable function. The* integral of f *is defined as,*

$$\int_{\mathbb{R}^n} f(x)\,dx = \sup_{0 \leq g \leq f} \int_{\mathbb{R}^n} g(x)\,dx$$

where g is a bounded measurable function supported on a finite measure set. As usual,

$$\int_E f(x)\,dx = \int_{\mathbb{R}^n} f(x)\chi_E(x)\,dx$$

since $f\chi_E \geq 0$ too, if $f \geq 0$.

Note that the supremum could be infinite and hence the integral could take infinite value.

Definition 3.3.2. *We say a non-negative function f is* Lebesgue integrable *if*

$$\int_{\mathbb{R}^n} f(x)\,dx < +\infty.$$

Exercise 58. Show the following properties of integral for non-negative measurable functions:

(i) (Linearity) For any two measurable functions f, g and $\alpha, \beta \in \mathbb{R}$,

$$\int_{\mathbb{R}^n} (\alpha f + \beta g)\,dx = \alpha \int_{\mathbb{R}^n} f\,dx + \beta \int_{\mathbb{R}^n} g\,dx.$$

(ii) (Additivity) For any two disjoint subsets $E, F \subset \mathbb{R}^n$ with finite measure

$$\int_{E\cup F} f\,dx = \int_E f\,dx + \int_F f\,dx.$$

(iii) (Monotonicity) If $f \leq g$, then

$$\int_{\mathbb{R}^n} f\,dx \leq \int_{\mathbb{R}^n} g\,dx.$$

In particular, if g is integrable and $0 \leq f \leq g$, then f is integrable.

(iv) If $f = g$ a.e. then $\int f = \int g$.

(v) If $f \geq 0$ and $\int f = 0$ then $f = 0$ a.e.

(vi) If f is integrable then f is finite a.e.

Do we have non-negative functions which are not Lebesgue integrable, i.e., for which the supremum is infinite?

Example 3.3. The function

$$f(x) = \begin{cases} \frac{1}{|x|} & \text{for } |x| \leq 1 \\ 0 & \text{for } |x| > 1. \end{cases}$$

is not integrable.

Following the definition of the notion of integral of a function, the immediate question we have been asking is the interchange of point-wise limit and integral. Thus, for a sequence of non-negative functions $\{f_k\}$ converging point-wise to f is

$$\int f = \lim_k \int f_k.$$

We have already seen in Example 3.2 that this is not true, in general. However, the following result is always true.

Lemma 3.3.3 (Fatou)**.** *Let $\{f_k\}$ be a sequence of non-negative measurable*[2] *functions converging point-wise a.e. to f, then*

$$\int f \leq \liminf_k \int f_k.$$

Proof. Let $0 \leq g \leq f$, where g is bounded with support on a set of finite measure E. Let $g_k(x) = \min(g(x), f_k(x))$, then g_k is measurable. Also, g_k is bounded by the bound of g and supported on E, since $g_k \leq g$ and g, f_k are non-negative for all k. We claim that g_k converges to g point-wise a.e. in E. Fix $x \in E$. Then either $g(x) = f(x)$ or $g(x) < f(x)$. Consider the case when $g(x) < f(x)$. For any given $\varepsilon > 0$ there is a $K \in \mathbb{N}$ such that $|f_k(x) - f(x)| < \varepsilon$ for all $k \geq K$. In particular, this is true for all $\varepsilon \leq f(x) - g(x)$. Thus, $g(x) \leq f(x) - \varepsilon < f_k(x)$ for all $k \geq K$ and hence $g_k(x) = g(x)$ for all $k \geq K$. On the other hand if $g(x) = f(x)$ then $g_k(x)$ is either $f(x)$ or $f_k(x)$ and will converge to $f(x)$. Thus, $g_k(x) \to g(x)$ a.e. and by the BCT $\int g_k \to \int g$. Moreover, $g_k \leq f_k$ and, by monotonicity, $\int g_k \leq \int f_k$ and therefore,

$$\int g = \lim_k \int g_k = \liminf_k \int g_k \leq \liminf_k \int f_k.$$

[2]Note that we do not demand integrability

Thus,

$$\int f = \sup \int g \leq \liminf_k \int f_k.$$

□

Exercise 59. Give an example of a situation where the we have strict inequality in Fatou's lemma.

Observe that Fatou's lemma basically says that the interchange of limit and integral is valid almost half way and what may go wrong is that

$$\limsup_k \int f_k > \int f.$$

Corollary 3.3.4. *Let $\{f_k\}$ be a sequence of non-negative measurable functions converging point-wise a.e. to f and $f_k(x) \leq f(x)$, then*

$$\int f = \lim_k \int f_k.$$

Proof. By monotonicity, $\int f_k \leq \int f$ and hence

$$\limsup_k \int f_k \leq \int f \leq \liminf_k \int f_k.$$

The second inequality is due to Fatou's lemma and hence we have

$$\int f = \lim_k \int f_k.$$

□

Corollary 3.3.5 (Monotone Convergence Theorem)**.** *Let $\{f_k\}$ be an increasing sequence of non-negative measurable functions converging point-wise a.e. to f, i.e., $f_k(x) \leq f_{k+1}(x)$, for all k, then*

$$\int f = \lim_k \int f_k.$$

Proof. Since f_k is increasing sequence, we have $f_k(x) \leq f(x)$ and hence we have our result by previous corollary. □

The MCT is not true for a decreasing sequence of functions. Consider $f_k = \chi_{[k,\infty)}$ on $\mathbb{R}$. f_k are non-negative and measurable functions on $\mathbb{R}$. The sequence f_k converges point-wise to $f \equiv 0$, since for each fixed $x \in \mathbb{R}$, $f_k(x) = 0$ for infinitely many k's. However, $\{f_k\}$ are decreasing, $f_{k+1}(x) \leq f_k(x)$ for all x and k. Moreover, $\int f_k = \infty$ and $\int f = 0$. Thus,

$$\int f \neq \lim_k \int f_k.$$

Corollary 3.3.6. *Let $\{f_k\}$ be a sequence of non-negative measurable functions. Then*

$$\int \left(\sum_{k=1}^{\infty} f_k(x) \right) dx = \sum_{k=1}^{\infty} \int f_k(x)\, dx.$$

Proof. Set $g_m(x) = \sum_{k=1}^{m} f_k(x)$ and $g(x) = \sum_{k=1}^{\infty} f_k(x)$. g_m are measurable and $g_m(x) \leq g_{m+1}(x)$ and g_m converges point-wise g. Thus, by MCT,

$$\int g = \lim_m \int g_m.$$

Thus,

$$\int \sum_{k=1}^{\infty} f_k(x) = \int g = \lim_m \int g_m = \lim_m \sum_{k=1}^{m} \int f_k(x) = \sum_{k=1}^{\infty} \int f_k(x).$$

□

The highlight of Fatou's lemma and its corollary is that they all remain true for a measurable function, i.e., we do allow the integrals to take ∞. We end this section by giving a different proof to the First Borel-Cantelli theorem proved in Theorem 2.3.13.

Theorem 3.3.7 (First Borel-Cantelli Lemma)**.** *If $\{E_i\}_1^{\infty} \subset \mathcal{L}(\mathbb{R}^n)$ be a countable collection of measurable subsets of $\mathbb{R}^n$ such that $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. Then $E := \cap_{k=1}^{\infty} \cup_{i=k}^{\infty} E_i$ has measure zero.*

Proof. Define $f_k := \chi_{E_k}$ and $f = \sum f_k$. Since f_k are non-negative, we have from the above corollary that

$$\int f = \sum_k \mu(E_k) < +\infty.$$

Thus, $f \in L^1(\mathbb{R}^n)$ and hence is finite a.e. Thus, the set $F := \{x \in \mathbb{R}^n \mid f(x) = \infty\}$ has measure zero. We claim that $E = F$. If $x \in F$, then $\sum_k \mu(E_k) = \infty$ implies that $x \in E_k$, for infinitely many k (a fact observed before) and hence $x \in E$. Conversely, if $x \in E$, then $x \in E_k$ for infinitely many k and hence $x \in F$. Thus, $\mu(E) = 0$. $\square$

3.4 General Integrable Functions

In this section, we try to extend our notion of integral to all other measurable functions. Recall that any function f can be decomposed in to $f = f^+ - f^-$ and both f^+, f^- are non-negative. Note that if f is measurable, both f^+ and f^- are measurable. We now give the definition of the integral of measurable functions.

Definition 3.4.1. *The* Lebesgue integral *of any measurable real-valued function f on $\mathbb{R}^n$ is defined as*

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

Any measurable function f is said to be Lebesgue integrable *if*

$$\int_{\mathbb{R}^n} |f(x)|\,dx < +\infty.$$

Any measurable function f is said to be locally Lebesgue integrable *if*

$$\int_K |f(x)|\,dx < +\infty,$$

for all compact subsets $K \subset \mathbb{R}^n$.

Observe that if f is measurable then $|f|$ is measurable. But the converse is not true (cf. Exercise 38). In view of this, one may have a non-measurable function f which is Lebesgue integrable. To avoid this situation, we assume the measurability of f in the definition of Lebesgue integrability of f.

Why use $|f|$ in the definition of integrability? For f to be integrable both f^+ and f^- should both be integrable which is true iff $|f|$ is integrable.

Exercise 60. Show that the definition of integral of f is independent of its decomposition $f = f_1 - f_2$ where $f_i \geq 0$ for $i = 1, 2$.

Exercise 61. The function $f : [-1, 0) \cup (0, 1] \to \mathbb{R}$ defined as $f(x) = \frac{1}{x}$ is not Lebesgue integrable although the improper integral (p.v.) exists.

Proof. The principal value intergal exists because

$$\lim_{\varepsilon \to 0} \int_{-1}^{-\varepsilon} \frac{1}{x}\, dx + \int_{\varepsilon}^{1} \frac{1}{x}\, dx = 0.$$

But the Lebesgue integrals are $\int_{-1}^{0} 1/x = -\infty$ and $\int_0^1 1/x = \infty$. □

Exercise 62. Consider $f(x) = \frac{\sin x}{x}$ on $[0, \infty)$. Using contour integration one can show that f is Riemann integrable (improper) and is equal to $\pi/2$. However, f is not Lebesgue integrable since $\int_0^\infty f^+ = \int_0^\infty f^- = \infty$.

Exercise 63. The function $f : (0, 1] \to \mathbb{R}$ defined as $f(x) = \frac{1}{x}\sin(\frac{1}{x}) + \cos(\frac{1}{x})$ is not Lebesgue integrable although the improper integral exists.

Proof. The improper intergal exists because

$$\lim_{\varepsilon \to 0} \int_{\varepsilon}^{1} \left[\frac{1}{x}\sin(\frac{1}{x}) + \cos(\frac{1}{x})\right] dx = \lim_{\varepsilon \to 0} \int_{\varepsilon}^{1} \frac{d}{dx}\left[x\cos(\frac{1}{x})\right] dx = \cos 1.$$

□

Definition 3.4.2. *A complex valued measurable function $f = u + iv$ on $\mathbb{R}^n$ is said to be integrable if*

$$\int_{\mathbb{R}^n} |f(x)|\, dx = \int_{\mathbb{R}^n} \left(u^2(x) + v^2(x)\right)^{1/2} dx < +\infty.$$

and the integral of f is given by

$$\int f = \int u + i \int v.$$

Exercise 64. Show that a complex-valued function is integrable iff both its real and imaginary parts are integrable.

Exercise 65. Show that all the properties of integral listed in Exercise 54 and Exercise 58 is also valid for a general integrable function.

The space of all real-valued measurable integrable functions on $\mathbb{R}^n$ is denoted by $L^1(\mathbb{R}^n)$. Thus, $L^1(\mathbb{R}^n) \subset M(\mathbb{R}^n)$. We will talk more on these spaces in the next section. We introduce this notation early in here only to use them in the statements of our results.

We highlight here that the non-negativity hypothesis in all results proved in the previous section (Fatou's lemma and its corollaries) can be replaced with a lower bound $g \in L^1(\mathbb{R}^n)$, since then we work with $g_k = f_k - g \geq 0$.

As usual we prove a result concerning the interchange of limit and integral, called the Dominated Convergence Theorem.

Theorem 3.4.3 (Dominated Convergence Theorem)**.** *Let $\{f_k\}$ be a sequence of measurable functions converging point-wise a.e. to f. If $|f_k(x)| \leq g(x)$, for all k, such that $g \in L^1(\mathbb{R}^n)$ then $f \in L^1(\mathbb{R}^n)$ and*

$$\int f = \lim_{k\to\infty} \int f_k.$$

Proof. Since $|f_k| \leq g$, $|f| \leq g$ and since $g \in L^1(\mathbb{R}^n)$, by monotonicity, $f \in L^1(\mathbb{R}^n)$. Note that $f_k(x) \leq |f_k(x)| \leq g(x)$. Hence, $g - f_k \geq 0$ and converges point-wise a.e. to $g - f$. By Fatou's lemma,

$$\int (g - f) \leq \liminf \int (g - f_k).$$

Therefore,

$$\int g - \int f \leq \int g + \liminf \left(- \int f_k \right) = \int g - \limsup \int f_k.$$

Thus, $\limsup \int f_k \leq \int f$, since $\int g$ is finite. Repeating above argument for the non-negative function $g + f_k$, we get $\int f \leq \liminf \int f_k$. Thus, $\int f = \lim \int f_k$. □

Exercise 66 (Generalised Dominated Convergence Theorem). Let $\{g_k\} \subset L^1(\mathbb{R}^n)$ be a sequence of integrable functions converging point-wise a.e. to $g \in L^1(\mathbb{R}^n)$. Let $\{f_k\}$ be a sequence of measurable functions converging point-wise a.e. to f and $|f_k(x)| \leq g_k(x)$. Then $f \in L^1(\mathbb{R}^n)$ and further if

$$\lim_k \int g_k = \int g$$

then

$$\lim_{k\to\infty} \int f_k = \int f.$$

Example 3.4. The Lebesgue dominated convergence theorem is a weaker statement than demanding uniform convergence. Consider $f_k(x) = x^k$ on $[0,1)$. $f_k(x) \to 0$ point-wise on $[0,1)$. However, the convergence is not uniform. But $|f_k| \leq 1$ and

$$\int_{[0,1]} x^k \, dx = \int_0^1 x^k \, dx = \frac{1}{k+1}$$

converges to zero.

Example 3.5. We have already seen using ψ_k in Example 3.2 that the bound by g in the hypothesis of DCT cannot be done away with. In fact, one can modify ψ_k in that example to have functions whose integrals diverge. For instance, choose $f_k(x) = k\psi_k = k^2\chi_{(0,1/k)}$ which point-wise converges to zero and $\int f_k = k$ which diverges.

Example 3.6. The condition that $g \in L^1(\mathbb{R}^n)$ is also crucial. For instance, let $f_k(x) = 1/k\chi_{[0,k]}$ and $|f_k(x)| \leq 1$ on $\mathbb{R}$. Note that f_k converge uniformly to zero, $\int f_k = 1$ do not converge to zero. Why? Because $g \equiv 1$ is not in $L^1(\mathbb{R})$.

Corollary 3.4.4. *Let $\{f_k\} \subset L^1(\mathbb{R}^n)$ such that*

$$\sum_{k=1}^{\infty} \int |f_k| \, dx < +\infty.$$

Then $\sum_{k=1}^{\infty} f_k(x) \in L^1(\mathbb{R}^n)$ and

$$\int \left(\sum_{k=1}^{\infty} f_k(x) \right) dx = \sum_{k=1}^{\infty} \int f_k(x) \, dx.$$

Proof. Let $g := \sum_{k=1}^{\infty} |f_k|$. Since $|f_k|$ is a non-negative sequence, by a corollary to Fatou's lemma, we have that

$$\int g(x) \, dx = \int \left(\sum_{k=1}^{\infty} |f_k| \right) dx = \sum_{k=1}^{\infty} \int |f_k| \, dx < +\infty.$$

Thus, $g \in L^1(\mathbb{R}^n)$. Now, consider

$$\left| \sum_{k=1}^{\infty} f_k(x) \right| \leq \sum_{k=1}^{\infty} |f_k(x)| = g.$$

Therefore $\sum_{k=1}^{\infty} f_k(x) \in L^1(\mathbb{R}^n)$, since $g \in L^1(\mathbb{R}^n)$. Consider the partial sum

$$F_m(x) = \sum_{k=1}^{m} f_k(x).$$

Note that $|F_m(x)| \leq g(x)$ for all k and $F_m(x) \to f(x)$ a.e.. Thus, by DCT, $\lim_m \int F_m = \int f$. By finite additivity of integals, we have

$$\int f = \lim_m \int F_m = \lim_m \sum_{k=1}^{m} \int f_k = \sum_{k=1}^{\infty} \int f_k.$$

Hence proved. □

Note that the BCT (cf. Theorem 3.2.3) had a stronger statement than DCT above. In fact, we can prove a similar statement for DCT.

Exercise 67. Let $\{f_k\}$ be a sequence of measurable functions converging point-wise a.e. to f. If $|f_k(x)| \leq g(x)$ such that $g \in L^1(\mathbb{R}^n)$ then

$$\lim_{k\to\infty} \int |f_k - f| = 0.$$

Proof. Note that $|f_k - f| \leq 2g$ and by DCT

$$\lim_k \int |f_k - f| = 0.$$

□

The above exercise could also be proved without using DCT and it is good enough proof to highlight here.

Theorem 3.4.5. *Let $\{f_k\}$ be a sequence of measurable functions converging point-wise a.e. to f. If $|f_k(x)| \leq g(x)$ such that $g \in L^1(\mathbb{R}^n)$ then*

$$\lim_{k\to\infty} \int |f_k - f| = 0.$$

In particular,

$$\lim_{k\to\infty} \int f_k = \int f.$$

Proof. Note that $|f(x)| \leq g(x)$, since f is point-wise limit of f_k. Let $E_k := \{x \mid x \in B_k(0) \text{ and } g(x) \leq k\}$. Note that g is non-negative. Set $g_k(x) := g(x)\chi_{E_k}(x)$ is measurable, integrable and non-negative. Also, $g_k(x) \leq g_{k+1}(x)$ and $g_k(x)$ converges point-wise to $g(x)$. By MCT, we have

$$\lim_k \int g_k = \int g.$$

Thus, for any given $\varepsilon > 0$, there exists a $K \in \mathbb{N}$ such that

$$\int_{E_k^c} g < \frac{\varepsilon}{4} \quad \forall k \geq K.$$

For the K obtained above, f_k restricted to E_K is uniformly bounded by K. Since $f_k(x) \to f(x)$ point-wise a.e. on E_K, by BCT, there exists a $K' \in \mathbb{N}$

$$\int_{E_K} |f_k - f| < \frac{\varepsilon}{2} \quad k \geq K'.$$

We have

$$\begin{aligned} \int |f_k - f| &= \int_{E_K} |f_k - f| + \int_{E_K^c} |f_k - f| \\ &\leq \int_{E_K} |f_k - f| + 2\int_{E_K^c} g \\ &< \frac{\varepsilon}{2} + 2\frac{\varepsilon}{4} = \varepsilon \quad \forall k \geq K'. \end{aligned}$$

Hence $\lim_k \int |f_k - f| \to 0$. □

The idea of the proof above actually suggests the following result.

Proposition 3.4.6. *Let $f \in L^1(\mathbb{R}^n)$. Then, for every given $\varepsilon > 0$,*

(i) There exists a ball $B \subset \mathbb{R}^n$ of finite measure such that

$$\int_{B^c} |f| < \varepsilon.$$

(ii) (Absolute Continuity) There exists a $\delta > 0$ such that

$$\int_E |f| < \varepsilon \quad \textit{whenever } \mu(E) < \delta.$$

Proof. Let $g(x) := |f(x)|$ and hence $g \geq 0$.

(i) Let $B_k := B_k(0)$ denote the ball of radius k centred at origin. Set $g_k(x) := g(x)\chi_{B_k}(x)$ is measurable and non-negative. Also, $g_k(x) \leq g_{k+1}(x)$ and $g_k(x)$ converges point-wise to $g(x)$. By MCT, we have

$$\lim_k \int g_k = \int g.$$

Thus, for the given $\varepsilon > 0$, there exists a $K \in \mathbb{N}$ such that $\int |f| - \int |f|\chi_{B_k} < \varepsilon$ for all $k \geq K$. Hence,

$$\varepsilon > \int (1 - \chi_{B_k})|f| = \int_{B_k^c} |f| \quad \forall k \geq K.$$

(ii) Let $E_k := \{x \mid g(x) \leq k\}$ and $g_k(x) = g(x)\chi_{E_k}(x)$. g_k is non-negative measurable function and $g_k(x) \leq g_{k+1}(x)$. Again, by MCT, there exists a $K \in \mathbb{N}$ such that $\int g - \int g_k < \varepsilon/2$ for all $k \geq K$. For any $E \in \mathcal{L}(\mathbb{R}^n)$,

$$\begin{aligned} \int_E g &= \int_E (g - g_K) + \int_E g_K \\ &\leq \int (g - g_K) + \int_E g_K \\ &\leq \int (g - g_K) + K\mu(E) \end{aligned}$$

Now, choose $\delta > 0$ such that $\delta < \frac{\varepsilon}{2K}$. If $\mu(E) < \delta$ then

$$\int_E g \leq \int (g - g_K) + K\mu(E) < \frac{\varepsilon}{2} + K\frac{\varepsilon}{2K} = \varepsilon.$$

□

The first part of the proposition above suggests that for integrable functions, the "integral of the function" vanishes as we approach infinity. However, this is not same as saying the function vanishes point-wise as $|x|$ approaches infinity.

Example 3.7. Consider the real-valued function f on $\mathbb{R}$

$$f(x) = \begin{cases} x & x \in \mathbb{Z} \\ 0 & x \in \mathbb{Z}^c. \end{cases}$$

$f = 0$ a.e. and $\int_{B^c} f = 0$, however $\lim_{x \to \infty} f(x) = +\infty$.

Can we have a continuous function in the above example?

Example 3.8. Let $f = \sum_{k=1}^{\infty} k\chi_{[k,k+1/k^3)}$. Note that

$$\sum_{k=1}^{\infty} \int k\chi_{\left[k,k+\frac{1}{k^3}\right)}\, dx = \sum_{k=1}^{\infty} \frac{1}{k^2} < +\infty.$$

The integral of f is

$$\int f = \sum_{k=1}^{\infty} 1/k^2$$

and is in $L^1(\mathbb{R}^n)$. But

$$\limsup_{x\to+\infty} f(x) = +\infty.$$

In fact, this is true for a continuous function in $L^1(\mathbb{R})$ (extend the f continuously to $\mathbb{R}$). However, for a uniformly continuous function in $L^1(\mathbb{R})$ we will have $\lim_{|x|\to\infty} f(x) = 0$.

Example 3.9. The integrability assumption, i.e., $f \in L^1(\mathbb{R}^n)$ is crucial the absolute continuity property (ii). Consider $f(x) = 1/x$ in $(0,1)$. Then for all $\delta > 0$ $\int_0^\delta |f|$ is not necessarily small. than can be large

The absolute continuity property of the integral proved in the Proposition above is precisely the continuity of the integral.

Exercise 68. Let $f \in L^1([a,b])$ and

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

Then F is continuous on $[a,b]$.

Proof. Let $x \in (a,b)$. Consider

$$|F(x) - F(y)| = \left|\int_x^y f(t)\, dt\right| \le \int_x^y |f(t)|\, dt.$$

Since $f \in L^1([a,b])$, by absolute continuity, for any given $\varepsilon > 0$ there is a $\delta > 0$ such that for all $y \in E = \{y \in [a,b] \mid |x-y| < \delta\}$, we have $|F(x) - F(y)| < \varepsilon$. $\square$

3.5 Order of Integration

Theorem 3.5.1 (Fubini). *Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be an integrable function. Then*

(i) $f^y : \mathbb{R}^m \to \mathbb{R}$ *is integrable defined as* $f^y(x) := f(x, y)$*, for a.e.* $y \in \mathbb{R}^n$ *and* $f^x : \mathbb{R}^n \to \mathbb{R}$ *is integrable defined as* $f^x(y) := f(x, y)$*, for a.e.* $x \in \mathbb{R}^m$.

(ii) $y \mapsto \int_{\mathbb{R}^m} f^y(x)\,dx$ *is integrable on* $\mathbb{R}^n$ *and* $x \mapsto \int_{\mathbb{R}^n} f^x(y)\,dy$ *is integrable on* $\mathbb{R}^m$.

(iii) $\int_{\mathbb{R}^n}\int_{\mathbb{R}^m} f^y(x)\,dx = \int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy = \int_{\mathbb{R}^m}\int_{\mathbb{R}^n} f^x(y)\,dy.$

Proof. First we observe that it is enough to prove the results for f^y and similar arguments are valid for f^x. Let $\mathcal{F}$ denote all integrable functions on $\mathbb{R}^{m+n}$ satisfying (i), (ii) and (iii). We have to show that every integrable functions belongs to $\mathcal{F}$.

Step 1 We first observe that $\mathcal{F}$ is closed under finite linear combinations. If $\{f_i\}$ is a finite collection in $\mathcal{F}$ and $\{A_i\}$ is the collection of zero measure sets such that f_i^y is integrable, for all $y \in A_i^c$, then $\cup_i A_i$ is of measure zero and in its complement f_i^y is integrable for all i. Thus, if f is any finite linear combination of $\{f_i\}$ then f^y is integrable in the complement of $\cup_i A_i$ and (ii) and (iii) follows from the linearity of integral.

Step 2 Let $\{f_k\}$ be an increasing (or decreasing) sequence of non-negative functions in $\mathcal{F}$ converging pointwise to f and let us assume that f is integrable. We claim $f \in \mathcal{F}$. By MCT, we have

$$\int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy = \lim_{k\to\infty}\int_{\mathbb{R}^{m+n}} f_k(x,y)\,dx\,dy.$$

Let $\{A_k\}$ be the collection of zero measure sets such that f_k^y is integrable, for all $y \in A_k^c$, then $\cup_k A_k$ is of measure zero and in its complement f_k^y is integrable for all k. By MCT,

$$g_k(y) := \int_{\mathbb{R}^m} f_k^y(x)\,dx$$

is an increasing sequence converging to

$$g(y) := \int_{\mathbb{R}^m} f^y(x)\,dx.$$

By MCT,

$$\int_{\mathbb{R}^n} g(y)\,dy = \lim_{k\to\infty} \int_{\mathbb{R}^n} g_k(y)\,dy.$$

But due to the assumption $f_k \in \mathcal{F}$ we know that

$$\int_{\mathbb{R}^n} g_k(y)\,dy = \int_{\mathbb{R}^{m+n}} f_k(x,y)\,dx\,dy.$$

Thus, we obtain

$$\int_{\mathbb{R}^n} g(y)\,dy = \int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy.$$

Since f is integrable, g is integrable and, hence, g is finite a.e. Thus, f^y is integrable for a.e. y and

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f^y(x)\,dx \right) dy = \int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy.$$

Hence, $f \in \mathcal{F}$.

Step 3 We now claim that $\chi_E \in \mathcal{F}$ where E is a measurable subset of $\mathbb{R}^{m+n}$ with finite measure.

(a) Suppose E is bounded open cell. Then $E = E_m \times E_n$ where E_m and E_n are cells of $\mathbb{R}^m$ and $\mathbb{R}^n$. Then χ_E^y is integrable for all y because

$$\chi_E^y = \begin{cases} \chi_{E_m} & y \in E_n \\ 0 & y \notin E_n. \end{cases}$$

and

$$\int_{\mathbb{R}^m} \chi_E^y(x)\,dx = \begin{cases} |E_m| & y \in E_n \\ 0 & y \notin E_n \end{cases} = |E_m|\chi_{E_n}$$

is also integrable. Therefore,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \chi_E^y(x)\,dx\,dy = |E_m||E_n| = |E| = \int_{\mathbb{R}^{m+n}} \chi_E\,dx\,dy$$

and $\chi_E \in \mathcal{F}$.

(b) Suppose E is a subset of the boundary of some closed cell. Then $\int_{\mathbb{R}^{m+n}} \chi_E(x,y)\,dx\,dy = 0$. Also, $\chi_E^y = 0$ a.e. in $\mathbb{R}^m$, for a.e. $y \in \mathbb{R}^n$, and

$$\int_{\mathbb{R}^m} \chi_E^y(x)\,dx = 0.$$

Hence,

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \chi_E^y(x)\,dx \right) dy = 0$$

and $\chi_E \in \mathcal{F}$.

(c) Suppose E is a finite almost disjoint union of closed cells, i.e. $E = \cup_{k=1}^N R_k$. Then χ_E is a linear combination of $\chi_{R_k^\circ}$, interior of R_k, and Γ_k, a subset of boundary of R_k. Thus, by Step 1, $\chi_E \in \mathcal{F}$.

(d) Suppose E is open with finite measure. Then $E = \cup_{k=1}^\infty R_k$ is countable almost disjoint union of closed cells. Then $f_k := \sum_{i=1}^k \chi_{R_i}$ increases to χ_E which is integrable ($\mu(E) < \infty$). Thus, by Step 2, $\chi_E \in \mathcal{F}$.

(e) Suppose E is a $\mathcal{G}_\delta$ finite measure subset of $\mathbb{R}^{m+n}$. The $E = \cap_{k=1}^\infty U_k$. Also, there exists an open set U such that $\mu(U) < \infty$ and $E \subset U$. Set $V_k := U \cap (\cap_{i=1}^k U_k)$. Then V_k is a decreasing sequence of open sets such that $E = \cap_{k=1}^\infty V_k$. Thus, χ_{V_k} is a decreasing sequence converging to χ_E. Since $\chi_{V_k} \in \mathcal{F}$, by Step 2, $\chi_E \in \mathcal{F}$.

(f) Suppose E is of zero measure. Then there exists a $\mathcal{G}_\delta$ set G such that $E \subset G$ and $\mu(G) = 0$. Since $\chi_G \in \mathcal{F}$, we have

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \chi_G^y(x)\,dx \right) dy = \int_{\mathbb{R}^{m+n}} \chi_G(x,y)\,dx\,dy = 0.$$

Therefore $\int_{\mathbb{R}^m} \chi_G^y(x)\,dx = 0$ for a.e. $y \in \mathbb{R}^n$. Thus $G^y := \{x \in \mathbb{R}^m \mid (x,y) \in G\}$ is of measure zero. Since $E^y \subset G^y$, we have $\mu(E^y) = 0$. Thus, $\int_{\mathbb{R}^m} \chi_E^y(x)\,dx = 0$ for a.e. $y \in \mathbb{R}^n$. Hence,

$$\int_{\mathbb{R}^{m+n}} \chi_E(x,y)\,dx\,dy = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \chi_E^y(x)\,dx \right) dy = 0.$$

(g) Suppose $E \subset \mathbb{R}^{m+n}$ is a measurable subset with a finite measure. Then there exists a finite measure $\mathcal{G}_\delta$ set G such that $E \subset G$ and $\mu(G \setminus E) = 0$. Consequently, $\chi_E = \chi_G - \chi_{G\setminus E}$, a finite linear combination of functions from $\mathcal{F}$. Thus, $\chi_E \in \mathcal{F}$.

Step 4 If $f \in L^1\mathbb{R}^{m+n}$ then $f = f^+ - f^-$ and an increasing sequence of simple functions converging to f^+ and f^-, respectively. Since simple functions are finite linear combinations. By step 3, $\chi_E \in \mathcal{F}$ and, by step 1, simple functions belong to $\mathcal{F}$. By step 2, f^+ and f^- are in $\mathcal{F}$ and, by step 1 again, $f \in \mathcal{F}$.

□

3.6 L^p Spaces

Recall that we already denoted, in the previous section, the class of integrable functions on $\mathbb{R}^n$ as $L^1(\mathbb{R}^n)$. What was the need for the superscript 1 in the notation?

Definition 3.6.1. *For any* $0 < p < \infty$*, a measurable function on* $E \in \mathcal{L}(\mathbb{R}^n)$ *is said to be* p-integrable *(Lebesgue) if*

$$\int_E |f(x)|^p \, dx < +\infty.$$

The space of all Lebesgue p-integrable functions on $E \in \mathcal{L}(\mathbb{R}^n)$ is denoted by $L^p(E)$.

In this sense, our integrable functions are precisely the 1-integrable functions.

Exercise 69. Show that $f \in L^p(E)$ then $|f|^p \in L^1(E)$.

The $p = \infty$ case is a generalisation of the uniform metric in the space of continuous bounded functions.

Definition 3.6.2. *A function f (not necessarily measurable) on $E \in \mathcal{L}(\mathbb{R}^n)$ is said to be* essentially bounded *if there exists a $0 < M < \infty$ such that $|f(x)| \le M$ a.e. in E, i.e., the set*

$$\{x \in \mathbb{R}^n \mid |f(x)| > M\}$$

has outer measure zero. The infimum of all such M is said to be the essential supremum *of f. The class of* measurable *essentially bounded function is denoted by $L^\infty(E)$.*

Exercise 70. Show that $L^p(E)$ forms a vector space over $\mathbb{R}$ (or $\mathbb{C}$) for $0 < p \leq \infty$.

Proof. The case $p = \infty$ is trivial. Consider the case $1 < p < \infty$. The closure under scalar multiplication is obvious. For closure under vector addition, we note that

$$|f(x) + g(x)| \leq |f(x)| + |g(x)|$$

and hence $|f(x) + g(x)|^p \leq (|f(x)| + |g(x)|)^p$ for all $p > 0$. Let $0 < p < \infty$ and $a, b \geq 0$. Assume wlog that $a \leq b$ (else we swap their roles). Thus, $a + b \leq 2b = 2\max(a, b)$ and therefore

$$(a + b)^p \leq 2^p b^p \leq 2^p(a^p + b^p).$$

Using this we get

$$(|f(x)| + |g(x)|)^p \leq 2^p|f(x)|^p + 2^p|g(x)|^p.$$

Therefore,

$$\int |f(x) + g(x)|^p \leq 2^p \int |f(x)|^p + 2^p \int |g(x)|^p < +\infty.$$

□

We introduce the notion of "length", called norm, on $L^p(E)$ for all $0 < p \leq \infty$.

Definition 3.6.3. *For all $0 < p < \infty$, we define the* norm *of $f \in L^p(E)$ as*

$$\|f\|_p := \left(\int_E |f(x)|^p \, dx\right)^{1/p},$$

which is finite. For any $f \in L^\infty(E)$, we define its norm *as*

$$\|f\|_\infty := \inf_M \{M \mid |f(x)| \leq M \text{ a.e.}\}.$$

Sometimes it is common to write $\|f\|_p$ as $\|f\|_{E,p}$ indicating the domain of f, however we shall restrain from complicating our notation and keep track of the domain of f wherever necessary. Let us now observe some properties of the norm. Observe that $\|f\|_p \geq 0$.

Exercise 71. Show that for each scalar $\lambda \in \mathbb{R}$ (or $\mathbb{C}$),

$$\|\lambda f\|_p = |\lambda| \|f\|_p \quad \forall f \in L^p(E).$$

(This is the reason for having the exponent $1/p$ in the definition of norm)

Note that the norm of zero function, $f \equiv 0$ is zero, but the converse is not true.

Exercise 72. For each $0 < p \leq \infty$, show that $\|f\|_p = 0$ iff $f = 0$ a.e.

Observe from the above exercise that the "length" we defined is short of being a "real length" (usually called semi-norm). In other words, we have non-zero vectors whose length is zero. To fix this issue, we inherit the equivalence relation of $M(\mathbb{R}^n)$ defined in Definition 2.4.5 to $L^p(\mathbb{R}^n)$. Thus, in the quotient space $L^p(E)/\sim$ length of all non-zero vectors is non-zero. In practice we always work with the quotient space $L^p(E)/\sim$ but write it as $L^p(E)$. Hence the remark following Definition 2.4.5 holds true for $L^p(E)$ (as the quotient space).

It now remains to show the triangle inequality of the norm. Proving triangle inequality is a problem due to the presence of the exponent $1/p$ (which was introduced for dilation property). For instance, the triangle inequality is true without the exponent $1/p$ in the definition of norm.

Exercise 73. Let $E \in \mathcal{L}(\mathbb{R}^n)$. Show that for $0 < p < 1$ and $f, g \in L^p(E)$ we have

$$\|f + g\|_p^p \leq \|f\|_p^p + \|g\|_p^p.$$

Proof. Let $0 < p < 1$ and $a, b \geq 0$. Assume wlog that $a \leq b$ (else swap their roles). For a fixed $p \in (0,1)$, the function x^p satisfies the hypotheses of MVT in $[b, a+b]$ and hence

$$(a+b)^p - b^p = pc^{p-1}a \quad \text{for some } c \in (b, a+b).$$

Now, since $p - 1 < 0$, we have

$$(a+b)^p = b^p + pc^{p-1}a \leq b^p + pb^{p-1}a \leq b^p + pa^{p-1}a = b^p + a^p.$$

Thus,

$$(a+b)^p \leq a^p + b^p.$$

Using this we have

$$(|f(x)| + |g(x)|)^p \leq |f(x)|^p + |g(x)|^p.$$

Therefore,

$$\int |f(x)+g(x)|^p \leq \int |f(x)|^p + \int |g(x)|^p < +\infty.$$

□

Note that we have not proved $\|f+g\|_p \leq \|f\|_p + \|g\|_p$, which is the triangle inequality. In fact, triangle inequality is false.

Example 3.10. Let $f = \chi_{[0,1/2)}$ and $g = \chi_{[1/2,1]}$ on $E = [0,1]$ and let $p = 1/2 < 1$. Note that $\|f+g\|_p = 1$ but $\|f\|_p = \|g\|_p = 2^{-(1/p)}$ and hence $\|f\|_p + \|g\|_p = 2^{1-1/p} < 1$. Thus, $\|f+g\|_p > \|f\|_p + \|g\|_p$.

Exercise 74. Show that for $0 < p < 1$ and $f, g \in L^p(E)$,

$$\|f+g\|_p \leq 2^{(1/p)-1}\left(\|f\|_p + \|g\|_p\right).$$

This is called the *quasi-triangle inequality.*

What is happening in reality is that the best constant for triangle inequality is $\max(2^{(1/p)-1}, 1)$, for all $p > 0$. Thus, when $p > 1$ the maximum is 1 and we have the triangle inequality of the norm for $p \geq 1$, called the Minkowski inequality. To prove this, we would need the general form of Cauchy-Schwarz inequality, called Hölder's inequality. For each $1 < p < \infty$, we associate with it a *conjugate* exponent q such that $1/p + 1/q = 1$. If $p = 1$ then we set $q = \infty$ and vice versa.

Theorem 3.6.4 (Hölder's Inequality)**.** *Let $E \in \mathcal{L}(\mathbb{R}^n)$ and $1 \leq p \leq \infty$. If $f \in L^p(E)$ and $g \in L^q(E)$, where the q is the conjugate exponent corresponding to p, then $fg \in L^1(E)$ and*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \tag{3.6.1}$$

Proof. If either f or g is a zero function a.e then the result is trivially true. Therefore, we assume wlog that both f and g have non-zero norm. Let $p = 1$ and $f \in L^1(E)$ and $g \in L^\infty(E)$. Consider

$$\int_E |fg| \leq \text{ess sup}_{x\in E}|g(x)| \int_E |f| = \|g\|_\infty \|f\|_1 < +\infty$$

Thus, $fg \in L^1(E)$. Let $1 < p < \infty$ and $f \in L^p(E)$ and $g \in L^q(E)$. If either $\|f\|_p = 0$ or $\|g\|_q = 0$, then equality holds trivially. Thus, we assume wlog that both $\|f\|_p, \|g\|_q > 0$. Set $f_1 = \frac{1}{\|f\|_p} f \in L^p(E)$ and $g_1 = \frac{1}{\|g\|_q} g \in L^q(E)$ with $\|f_1\|_p = \|g_1\|_q = 1$. Recall the AM-GM inequality (cf. (**??**)),

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Using this we get

$$\begin{aligned} |f_1(x)g_1(x)| &\leq \frac{1}{p}|f_1(x)|^p + \frac{1}{q}|g_1(x)|^q \\ \frac{1}{\|f\|_p}|f(x)|\frac{1}{\|g\|_q}|g(x)| &\leq \frac{1}{p\|f\|_p^p}|f(x)|^p + \frac{1}{q\|g\|_q^q}|g(x)|^q. \end{aligned}$$

Now, integrating both sides w.r.t the Lebesgue measure, we get

$$\begin{aligned} \int |fg| &\leq \|f\|_p\|g\|_q \left(\frac{1}{p\|f\|_p^p}\|f\|_p^p + \frac{1}{q\|g\|_q^q}\|g\|_q^q\right) \\ &= \|f\|_p\|g\|_q. \end{aligned}$$

Hence $fg \in L^1(E)$. □

Remark 3.6.5. Equality holds in (3.6.1) iff equality holds in (**??**) which happens, by Theorem **??**, iff $\frac{|f(x)|^p}{\|f\|_p^p} = \frac{|g(x)|^q}{\|g\|_q^q}$, for a.e. $x \in E$. Thus, $|f(x)|^p = \lambda |g(x)|^q$ where $\lambda = \frac{\|f\|_p^p}{\|g\|_q^q}$.

Exercise 75. Show that for $0 < p < 1$, $f \in L^p(E)$ and $g \in L^q(E)$ where the the conjugate exponent of p (now it is negative),

$$\|fg\|_1 \geq \|f\|_p\|g\|_q.$$

Theorem 3.6.6 (Minkowski Inequality)**.** *Let $E \in \mathcal{L}(\mathbb{R}^n)$ and $1 \leq p \leq \infty$. If $f, g \in L^p(E)$ then $f + g \in L^p(E)$ and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \tag{3.6.2}$$

Proof. The proof is obvious for $p = 1$, since $|f(x) + g(x)| \leq |f(x)| + |g(x)|$. Let $1 < p < \infty$ and q be the conjugate exponent of p. Observe that

$$\begin{aligned}
\int |f(x)+g(x)|^p &= \int |f(x)+g(x)|^{p-1}|f(x)+g(x)| \\
&\leq \int |f(x)+g(x)|^{p-1}|f(x)| + \int |f(x)+g(x)|^{p-1}|g(x)| \\
&\leq \left\|(f+g)^{p-1}\right\|_q \|f\|_p + \left\|(f+g)^{p-1}\right\|_q \|g\|_p \\
&\leq \left\|(f+g)^{p-1}\right\|_q (\|f\|_p + \|g\|_p) \\
&= \left(\int |f(x)+g(x)|^p\right)^{1/q} (\|f\|_p + \|g\|_p) \\
\|f+g\|_p^p &= \|f+g\|_p^{p/q} (\|f\|_p + \|g\|_p) \\
\|f+g\|_p &= \|f\|_p + \|g\|_p.
\end{aligned}$$

Hence $f + g \in L^p(E)$. □

Exercise 76. Show that for $0 < p < 1$ and $f, g \in L^p(E)$ such that f, g are non-negative

$$\|f+g\|_p \geq \|f\|_p + \|g\|_p$$

The triangle inequality fails for $0 < p < 1$ due to the presence of the exponent $1/p$ in the definition of $\|f\|_p$. Thus, for $0 < p < 1$, we also have the option of ignoring the $1/p$ exponent while defining $\|f\|_p$. Define the metric $d_p : L^p(\mathbb{R}^n) \times L^p(\mathbb{R}^n) \to [0, \infty)$ on $L^P(\mathbb{R}^n)$ such that $d_p(f, g) = \|f - g\|_p$ for $1 \leq p \leq \infty$ and $d_p(f, g) = \|f - g\|_p^p$ for $0 < p < 1$.

Exercise 77. Show that d_p is a metric on $L^p(\mathbb{R}^n)$ for $p > 0$.

Definition 3.6.7. *Let $E \in \mathcal{L}(\mathbb{R}^n)$. We say a sequence $\{f_k\}$ converges to f in $L^p(E)$, $p > 0$, if $d_p(f_k, f) \to 0$ as $k \to \infty$.*

Exercise 78. Show that if $f_k \to f$ in L^p then $f_k \xrightarrow{\mu} f$ in measure.

The converse is not true, in general. Moreover, L^p convergence does not imply almost uniform or point-wise a.e. However, they are true for a subsequence.

Exercise 79. If f_k converges to f in $L^p(E)$ then there exists a subsequence $\{f_{k_l}\}$ of $\{f_k\}$ such that $f_{k_l}(x) \to f(x)$ point-wise for a.e. $x \in E$.

Theorem 3.6.8 (Riesz-Fischer)**.** *For $p > 0$, $L^p(\mathbb{R}^n)$ is a complete metric space.*

In general, we ignore studying L^p for $0 < p < 1$ because its dual[3] is trivial vector space. Thus, henceforth we restrict ourselves to $1 \leq p \leq \infty$.

Exercise 80. Let $1 \leq p < \infty$. For what values of $\delta \in \mathbb{R}$ does $|x|^\delta \in L^p(B_1(0))$ where $B_1(0)$ is the unit ball of $\mathbb{R}^n$.

Proof. Consider

$$\int_{B_1(0)} |x|^{\delta p}\, dx = \int_{S_1(0)} \int_0^1 r^{\delta p + n - 1}\, dr\, d\sigma.$$

Thus, for $\delta p + n - 1 > -1$ or $\delta > \frac{-n}{p}$, the integral is finite and is equal to $\frac{\omega_n}{\delta p + n}$, where ω_n is the surface measure of the unit ball. Also, note that, for all $\frac{-n}{p} < \delta < 0$, $|x|^\delta$ has a blow-up near 0. □

Proposition 3.6.9. *Let $E \in \mathcal{L}(\mathbb{R}^n)$ be such that $\mu(E) < +\infty$. If $1 \leq p < r < \infty$ then $L^r(E) \subset L^p(E)$ and*

$$\|f\|_p \leq \frac{\mu(E)^{1/p}}{\mu(E)^{1/r}} \|f\|_r.$$

Proof. Let $f \in L^r(E)$. We need to show that $f \in L^p(E)$. Set $F = |f|^p$ and $G = 1$. Note that $F \in L^{r/p}(E)$ since

$$\int |F|^{r/p} = \int |f|^r < +\infty.$$

Applying Hölder's inequality, we get

$$\begin{aligned} \|f\|_p^p = \|FG\|_1 &\leq \|F\|_{r/p} \|G\|_{r/(r-p)} \\ &= \|f\|_r^p \mu(E)^{(r-p)/r} \\ \|f\|_p &\leq \mu(E)^{1/p - 1/r} \|f\|_r. \end{aligned}$$

□

Exercise 81. The inclusion obtained above is, in general, strict. For instance, observe from Exercise 80 that $1/\sqrt{x}$ on $[0, 1]$ is in $L^1([0, 1])$ but is not in $L^2([0, 1])$.

[3] will be introduced in a course of functional analysis

Proposition 3.6.10. *Let $E \in \mathcal{L}(\mathbb{R}^n)$ be such that $\mu(E) < +\infty$. If $f \in L^\infty(E)$ then $f \in L^p(E)$ for all $1 \leq p < \infty$ and*

$$\|f\|_p \to \|f\|_\infty \text{ as } p \to \infty.$$

Proof. Given $f \in L^\infty(E)$, note that

$$\|f\|_p = \left(\int_E |f|^p\right)^{1/p} \leq \|f\|_\infty \mu(E)^{1/p}.$$

Since $\mu(E)^{1/p} \to 1$ as $p \to \infty$, we get $\limsup_{p\to\infty} \|f\|_p \leq \|f\|_\infty$. Recall that $\|f\|_\infty$ is the infimum over all essential bounds of f. Thus, for every $\varepsilon > 0$ there is a $\delta > 0$ such that $\mu(E_\delta) \geq \delta$ where

$$E_\delta := \{x \in E \mid |f(x)| \geq \|f\|_\infty - \varepsilon\}.$$

Thus,

$$\|f\|_p^p \geq \int_{E_\delta} |f|^p \geq (\|f\|_\infty - \varepsilon)^p \, \delta.$$

Since $\delta^{1/p} \to 1$ as $p \to \infty$, we get $\liminf_{p\to\infty} \|f\|_p \geq \|f\|_\infty - \varepsilon$. Since choice of ε is arbitrary, we have $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$. □

Exercise 82. Let $\{f_k\}$ be a sequence of functions in $L^\infty(\mathbb{R}^n)$. Show that $\|f_k - f\|_\infty \to 0$ iff there is a set $E \in \mathcal{L}(\mathbb{R}^n)$ such that $\mu(E) = 0$ and $f_k \to f$ uniformly on E^c.

Proof. It is enough to show the result for $f \equiv 0$. Let $\|f_k\|_\infty \to 0$. Let $E_k := \{x \in \mathbb{R}^n \mid |f_k(x)| > \|f_k\|_\infty\}$. Note that $\mu(E_k) = 0$. Set $E = \cup_{k=1}^\infty E_k$. By sub-additivity of Lebesgue measure, $\mu(E) = 0$. Fix $\varepsilon > 0$. Then there is a $K \in \mathbb{N}$ such that for all $k \geq K$, $\|f_k\|_\infty < \varepsilon$. Choose any $x \in E^c$. Then $|f_k(x)| \leq \|f_k\|_\infty$ for all k. Thus, for all $x \in E^c$ and $k \geq K$, $|f_k(x)| < \varepsilon$.. Thus, $f_k \to 0$ uniformly on E^c.

Conversely, let $E \in \mathcal{L}(\mathbb{R}^n)$ be such that $\mu(E) = 0$ and $f_k \to 0$ uniformly on E^c. Fix $\varepsilon > 0$. For any $x \in E^c$, there exists a $K \in \mathbb{N}$ (independent of x) such that $|f_k(x)| < \varepsilon$ for all $k \geq K$. For each $k \geq K$,

$$\|f_k\|_\infty = \text{ess sup}_{x\in\mathbb{R}^n} |f_k(x)| = \sup_{x\in E^c} |f_k(x)| < \varepsilon.$$

Hence $\|f_k\|_\infty \to 0$. □

We now prove some density results of L^p spaces. Recall that in Theorem 2.4.11 and Theorem 2.4.12 we proved the density of simple function in $M(\mathbb{R}^n)$ under point-wise a.e. convergence. We shall now prove the density of simple functions in L^p spaces. We say a collection of functions $\mathcal{A} \subset L^p$ is dense in L^p if for every $f \in L^p$ and $\varepsilon > 0$ there is a $g \in \mathcal{A}$ such that $\|f - g\|_p < \varepsilon$.

Theorem 3.6.11. *Let $E \in \mathcal{L}(\mathbb{R}^n)$. The class of all simple [4] functions are dense in $L^p(E)$ for $1 \leq p < \infty$.*

Proof. Fix $1 \leq p < \infty$ and let $f \in L^p(E)$ such that $f \geq 0$. By Theorem 2.4.11 we have an increasing sequence of non-negative simple functions $\{\phi_k\}$ that converge point-wise a.e. to f and $\phi_k \leq f$ for all k. Thus,

$$|\phi_k(x) - f(x)|^p \leq 2^p |f(x)|^p$$

and by DCT we have

$$\lim_{k\to\infty} \|\phi_k - f\|_p^p = \lim_{k\to\infty} \int_E |\phi_k - f|^p \to 0.$$

For an arbitrary $f \in L^p(E)$, we use the decomposition $f = f^+ - f^-$ where $f^+, f^- \geq 0$. Thus we have sequences of simple functions $\{\phi_k\}$ and $\{\psi_k\}$ such that $\phi_m - \psi_m \to f$ in $L^p(E)$ (using triangle inequality). Thus, the space of simple functions is dense in $L^p(E)$. ☐

Theorem 3.6.12. *Let $E \in \mathcal{L}(\mathbb{R}^n)$. The space of all compactly supported continuous functions on E, denoted as $C_c(E)$ is dense in $L^p(E)$ for $1 \leq p < \infty$.*

Proof. It is enough to prove the result for a characteristic function χ_F, where $F \subset E$ such that F is bounded. By outer regularity, for a given $\varepsilon > 0$ there

[4] By our definition, simple function is non-zero on a finite measure. A simple function ϕ is a non-zero function on $\mathbb{R}^n$ having the (canonical) form

$$\phi(x) = \sum_{i=1}^{k} a_i 1_{E_i}$$

with disjoint measurable subsets $E_i \subset \mathbb{R}^n$ with $\mu(E_i) < +\infty$ and $a_i \neq 0$, for all i, and $a_i \neq a_j$ for $i \neq j$.

is an open (bounded) set Ω such that $\Omega \supset F$ and $\mu(\Omega \setminus F) < \varepsilon/2$. Also, by inner regularity, there is a compact set $K \subset F$ such that $\mu(F \setminus K) < \varepsilon/2$. By Urysohn lemma there is a continuous function $g : E \to \mathbb{R}$ such that $g \equiv 0$ on $E \setminus \Omega$, $g \equiv 1$ on K and $0 \leq g \leq 1$ on $\Omega \setminus K$. Note that $g \in C_c(E)$. Therefore,

$$\|\chi_F - g\|_p^p = \int_E |\chi_F - g|^p = \int_{\Omega\setminus K} |\chi_F - g|^p \leq \mu(\Omega \setminus K) = \varepsilon.$$

□

Aliter. Let $f \in L^p(E)$ and fix $\varepsilon > 0$. By Theorem 3.6.11, there is a simple function ϕ such that $\|\phi - f\|_p < \varepsilon/2$. Note that ϕ is supported on a finite measure set, by definition of simple funciton. Let $F := \operatorname{supp}(\phi)$ and $F \subset E$. By Luzin's theorem, there is a closed subset $\Gamma \subset F$ such that $\phi \in C(\Gamma)$ and

$$\mu(F \setminus \Gamma) < \left(\frac{\varepsilon}{2\|\phi\|_\infty}\right)^p.$$

Γ being a closed subset of finite measure set F, Γ is compact in E. Thus, we put ϕ to be zero on $\Gamma^c := E \setminus \Gamma$, call it g, and $g \in C_c(E)$ with $\operatorname{supp}(g) = \Gamma$. Further, by our construction, we have $|g(x)| \leq \|\phi\|_\infty$. Hence,

$$\|g - \phi\|_p = \|\phi\|_{p,\Gamma^c} = \|\phi\|_{p,F\setminus\Gamma} < \frac{\varepsilon}{2\|\phi\|_\infty}\|\phi\|_\infty = \frac{\varepsilon}{2}.$$

Therefore, $\|g - f\|_p < \varepsilon$. Thus, $C_c(E)$ is dense in $L^p(E)$. □

Example 3.11. The class of all simple functions is not dense in $L^\infty(E)$. The space $C_c(E)$ is not dense in $L^\infty(E)$, but is dense $C_0(E)$ with uniform norm, the space of all continuous function vanishing at infinity.

Theorem 3.6.13. *For $1 \leq p < \infty$, $L^p(\mathbb{R}^n)$ is separable but $L^\infty(\mathbb{R}^n)$ is not separable.*

3.7 Invariance of Lebesgue Integral

Recall that in section 15, we noted the invariance properties of the Lebesgue measure. In this section, we shall note the invariance properties of Lebesgue integral.

Definition 3.7.1. *For any function f on $\mathbb{R}^n$ we define its* translation *by a vector $y \in \mathbb{R}^n$, denoted $\tau_y f$, as*

$$\tau_y f(x) = f(x - y).$$

Similarly, one can define notion similar to reflection

$$\check{f}(x) = f(-x).$$

Also, dilation by $\lambda > 0$, is $f(\lambda x)$.

Exercise 83. Show that if $f \in L^1(\mathbb{R}^n)$

(i) (**Translation invariance**) then $\tau_y f \in L^1(\mathbb{R}^n)$, for every $y \in \mathbb{R}^n$, and $\int f = \int \tau_y f$.

(ii) (**Reflection**) then $\check{f} \in L^1(\mathbb{R}^n)$ and $\int f = \int \check{f}$.

(iii) (**Dilation**) and $\lambda > 0$, then $f(\lambda x) \in L^1(\mathbb{R}^n)$ and $\int f = \lambda^n \int f(\lambda x)$.

Chapter 4

Duality of Differentiation and Integration

The aim of this chapter is to identify the general class functions (within the framework of concepts developed in previous chapters) for which following is true:

1. (Derivative of an integral)

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

2. (Integral of a derivative)

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

We shall attempt to answer these questions in one-dimensional case to keep our attempt simple. In fact answering both these questions have far-reaching consequences not highlighted in this chapter. Let $f : [a, b] \to \mathbb{R}$ be Lebegue integrable and define

$$F(x) = \int_a^x f(t)\,dt \quad x \in [a, b].$$

Answering first question is equivalent to saying $F'(x) = f(x)$. Note that for a non-negative f, F is a monotonically increasing function. This observation motivates the study of monotone functions in the next section.

4.1 Monotone Functions

Recall that a function is said to be monotone if it preserves a given order. A function f is said to be monotonically increasing if $f(x) \leq f(y)$ whenever $x \leq y$. If $f(x) < f(y)$ whenever $x \leq y$, we say f is strictly increasing. We prove in this section that a monotone increasing function is differentiable a.e. A major tool for proving this result is the Vitali covering lemma.

Definition 4.1.1. *Let $E \subset \mathbb{R}^n$. We say a collection of balls $\mathcal{V}$ is a* Vitali covering *of E if for every $\varepsilon > 0$ and for every $x \in E$, there exist a ball $B \in \mathcal{V}$ such that $x \in B$ and $\mu(B) < \varepsilon$.*

In the definition above we allow the balls to be open or closed but do not allow degenerate balls consisting of single a point or lower dimensional balls. We now prove the Vitali's covering lemma which claims that one can extract a finite disjoint sub-cover of the Vitali cover such that it "almost" covers E. The proof is constructive.

Lemma 4.1.2 (Vitali Covering Lemma)**.** *Let $E \subset \mathbb{R}^n$ be an arbitrary subset such that $\mu^\star(E) < +\infty$ and let $\mathcal{V}$ be a Vitali covering of E. Then, for every $\varepsilon > 0$, there is a finite disjoint sub-collection $\{B_i\}_1^k \subset \mathcal{V}$ such that*

$$\mu^\star\left(E \setminus \cup_{i=1}^k B_i\right) < \varepsilon.$$

Proof. Without loss of generality, we assume that each ball in $\mathcal{V}$ is closed because

$$\mu^\star\left(E \setminus \cup_{i=1}^k B_i\right) = \mu^\star\left(E \setminus \cup_{i=1}^k \overline{B}_i\right).$$

Let Ω be an open subset of $\mathbb{R}^n$ such that $\mu(\Omega) < +\infty$ and $E \subset \Omega$. We now assume without loss of generality that $B \subseteq \Omega$ for all $B \in \mathcal{V}$. This is possible at the cost of throwing away some elements of the covering $\mathcal{V}$ because $\mu^\star(E) < +\infty$ and balls in $\mathcal{V}$ as small as possible.

First we pick a ball $B_1 \in \mathcal{V}$. If $E \subset B_1$ then we are done. Else, for each $k \geq 2$, we pick a ball B_k such that $B_k \cap (\cup_{i=1}^{k-1} B_i) = \emptyset$ and $\mu(B_k) > r_k/2$, where

$$r_k = \sup_{B \in \mathcal{V}} \{\mu(B) \mid B \cap \cup_{i=1}^{k-1} B_i = \emptyset\}.$$

Note that r_k is finite for all k, since $r_k \leq \mu(\Omega) < +\infty$. If the set over which the supremum is taken is an empty collection for some k, i.e., there is no $B \in \mathcal{V}$ such that $B \cap \cup_{i=1}^{k-1} B_i = \emptyset$, then we already have $E \subset \cup_{i=1}^{k-1} B_i$ and we are done. Otherwise, we have a countable disjoint collection of closed balls B_k such that $\mu(B_k) > r_k/2$. Also $\cup_k B_k \subset \Omega$ and hence, by monotonicity and additivity of μ, we get

$$+\infty > \mu(\Omega) \geq \mu(\cup_{k=1}^\infty B_k) = \sum_{k=1}^\infty \mu(B_k).$$

Thus, for any given $\varepsilon > 0$, there is $K \in \mathbb{N}$ such that

$$\sum_{k=K+1}^\infty \mu(B_k) < \frac{\varepsilon}{5^n}.$$

We wish to prove that $\mu^\star(E \setminus \cup_{i=1}^K B_i) < \varepsilon$. Let $x \in E \setminus \cup_{i=1}^K B_i$. Such an x exists because otherwise we would not have countable collection and our process would have stopped at Kth stage. For the chosen x we have a ball $B_x \in \mathcal{V}$ such that $B_x \cap \cup_{i=1}^K B_i = \emptyset$. Note that $\mu(B_x) \leq r_K$. We now claim that B_x intersects B_i for some $i > K$. Suppose, for every $i > K$, $B_x \cap B_i = \emptyset$, then $\mu(B_x) \leq r_i \to 0$ as i becomes large. Thus, B_x intersects B_i for some $i > K$. Let l be the smallest $i > K$ (or first instance) when B_x meets B_l. Hence $\mu(B_x) \leq r_l < 2\mu(B_l)$. Thus, $r_x < 2r_l$. We claim that $B_x \subset 5B_l$. Let $y \in B_x \cap B_l$. Then, $|x - y| \leq 2r_x$, where r_x is the radius of B_x. Also, $|y - x_l| \leq r_l$ where x_l is the centre of B_l and r_l is its radius. Thus,

$$|x - x_l| \leq |x - y| + |y - x_l| \leq 2r_x + r_l < 4r_l + r_l = 5r_l.$$

Hence $x \in 5B_l$ and $B_x \subset 5B_l$. Therefore,

$$E \setminus \cup_{i=1}^K B_i \subset \cup_{i=K+1}^\infty 5B_i$$

and

$$\mu^\star(E \setminus \cup_{i=1}^K B_i) \leq \mu(\cup_{i=K+1}^\infty 5B_i) \leq 5^n \sum_{i=K+1}^\infty \mu(B_i) < \varepsilon.$$

□

Lemma 4.1.3 (Riesz's Rising Sun Lemma)**.** *Let $g : [a, b] \to \mathbb{R}$ be a continuous function. If*

$$E := \{x \in (a, b) \mid g(x + h) > g(x) \text{ for some } h > 0\}$$

then E is either empty or open. In the latter case (E open) E is a countable union of disjoint intervals (a_k, b_k) with $g(a_k) = g(b_k)$ when $a \neq a_k$. For $a = a_k$, $g(a) \leq g(b_k)$.

Proof. Note that E is empty iff g is non-increasing. Thus, for a generic continuous function g, E is non-empty. Note that $E = \cup_{y\in(a,b)} E_y$ where $E_y := \{x \in (a, y) \mid g(x) < g(y)\}$. By the continuity of g, E_y is open and, hence, E is open. Thus, E can be written as a disjoint union of countably many open intervals (a_k, b_k). Since $a_k \notin E$, we have $g(a_k) \geq g(b_k)$. Suppose $g(a_k) > g(b_k)$, then by the continutiy of g, there is a $c \in (a_k, b_k)$ such that

$$g(c) = \frac{g(a_k) + g(b_k)}{2}.$$

Among all possible such c choose the one that is closest to b_k. Picking such a closest c to b_k is possible. If not, then the accumulation point of such c's should be b_k which is not possible because $g(b_k) \neq \frac{g(a_k)+g(b_k)}{2}$. Note that $c \in E$ and, hence, there is a $d > c$ such that $g(d) > g(c)$. Since $b_k \notin E$, we have $g(b_k) \geq g(x)$ for all $x \geq b_k$. Since $g(d) > g(c) > g(b_k)$, we have $d < b_k$. Also, since $g(d) > g(c) > g(b_k)$, by continuity of g, there is a $c_1 \in (d, b_k)$ such that $g(c_1) = g(c)$. Thus, c_1 contradicts the proximity of c with b_k. Hence, the hypothesis $g(a_k) > g(b_k)$ is false and, thus, $g(a_k) = g(b_k)$. □

Theorem 4.1.4. *Let $f : [a, b] \to \mathbb{R}$ be a monotone function. Then f has at most countably many discontinuity points. Conversely, given any countable set $E \subset \mathbb{R}$, there exists a monotone function $f : \mathbb{R} \to \mathbb{R}$ whose set of discontinuity is exactly E.*

Proof. Without loss of generality let us assume f is increasing. For each $x \in (a, b)$, define the left and right limit ($h > 0$)

$$f_+(x) := \lim_{h\to 0} f(x + h) \text{ and } f_-(x) := \lim_{h\to 0} f(x - h).$$

Then, $J(x) := f_+(x) - f_-(x) \geq 0$ is the jump of f at x. Thus, f is continuous at x iff $J(x) = 0$. For each $n \in \mathbb{N}$, define

$$E_n := \{x \in (a, b) \mid J(x) \geq \frac{1}{n}\}.$$

Note that the set of discontinuity points of f is precisely $\cup_{n=1}^{\infty} E_n$. Let $I : \{x_1, x_2, \ldots, x_k\}$ be a finite subset of E_n such that $x_1 < x_2 < \ldots < x_k$. Since f is increasing we have

$$f(a) \leq f_-(x_1) \leq f_+(x_1) \leq \ldots \leq f_-(x_k) \leq f_+(x_k) \leq f(b)$$

and

$$f(b) - f(a) \geq \sum_{i=1}^{k} [f_+(x_i) - f_-(x_i)] = \sum_{x \in I} J(x) \geq \frac{k}{n}.$$

Thus, the cardinality of E_n is at most the integer part of $n[f(b) - f(a)]$.

Conversely, let E be a countable set. If E is finite then construct a monotone linear function in the interval between two discontinuity points. Suppose $E = \{x_n\}$ is countable. For each $n \in \mathbb{N}$, define an increasing function $f_n : \mathbb{R} \to \mathbb{R}$ by

$$f_n(x) := \begin{cases} -\frac{1}{n^2} & \text{if } x < x_n \\ \frac{1}{n^2} & \text{if } x \geq x_n. \end{cases}$$

Note that f_n is discontinuous only at x_n. Define, for all $x \in \mathbb{R}$,

$$f(x) := \sum_{n=1}^{\infty} f_n(x).$$

Since $|f_n(x)| \leq \frac{1}{n^2}$ for all $x \in \mathbb{R}$. The series is uniformly convergent and, hence, f is well-defined and continuous at every point on which each f_n is continuous. Thus, f is continuous on $\mathbb{R} \setminus E$. We now prove that f is discontinuous at each point of E. Note that, for each $n \in \mathbb{N}$,

$$f = f_n + \sum_{i \neq n} f_i.$$

Since $\sum_{i \neq n} f_i$ is continuous at x_n and f_n is not continuous at x_n, f is discontinuous at x_n. Further, f is increasing because it is the pointwise limit of a sequence of increasing functions. $\square$

Example 4.1. There exists an increasing function $f : \mathbb{R} \to \mathbb{R}$ that is continuous at all irrational points and discontinuous at all rational points.

To prove the main result of this section, i.e., every increasing function is differentiable a.e., we need the following derivatives, called *Dini* derivatives.

Definition 4.1.5. *For any given function $f : \mathbb{R} \to \mathbb{R}$, we define the four*

Dini derivatives *of f at x as follows:*

$$\begin{aligned}
D^+f(x) &= \limsup_{h\to 0+} \frac{f(x+h)-f(x)}{h}, \\
D_+f(x) &= \liminf_{h\to 0+} \frac{f(x+h)-f(x)}{h}, \\
D^-f(x) &= \limsup_{h\to 0-} \frac{f(x+h)-f(x)}{h} = \limsup_{h\to 0+} \frac{f(x)-f(x-h)}{h}, \\
D_-f(x) &= \liminf_{h\to 0-} \frac{f(x+h)-f(x)}{h} = \liminf_{h\to 0+} \frac{f(x)-f(x-h)}{h}.
\end{aligned}$$

Note that these numbers always exist and could be infinity. Also, we always have $D^+f(x) \geq D_+f(x)$ and $D^-f(x) \geq D_-f(x)$. If $D^+f(x) = D_+f(x) \neq \pm\infty$ then we say the right-hand derivative of f exists at x. Similarly, if $D^-f(x) = D_-f(x) \neq \pm\infty$ then we say the left-hand derivative of f exists at x. We say f is differentiable at x if $D^+f(x) = D_+f(x) = D^-f(x) = D_-f(x) \neq \pm\infty$ and $f'(x) = D^+f(x)$. Observe that for a increasing function the Dini derivatives are all non-negative.

Exercise 84. Show that, for a given $f : \mathbb{R} \to \mathbb{R}$, if $g(x) = -f(-x)$ for all $x \in \mathbb{R}$ then $D^+g(x) = D^-f(-x)$ and $D_-g(x) = D_+f(-x)$.

Proof. Consider

$$\begin{aligned}
D^+g(x) &= \limsup_{h\to 0+} \frac{g(x+h)-g(x)}{h} \\
&= \limsup_{h\to 0+} \frac{-f(-x-h)+f(-x)}{h} = D^-f(-x).
\end{aligned}$$

Similarly,

$$\begin{aligned}
D_-g(x) &= \liminf_{h\to 0+} \frac{g(x)-g(x-h)}{h} \\
&= \liminf_{h\to 0+} \frac{-f(-x)+f(-x+h)}{h} = D_+f(-x).
\end{aligned}$$

□

Theorem 4.1.6. *If $f : [a,b] \to \mathbb{R}$ is a increasing function then f is differentiable a.e. in $[a,b]$. Moreover, the derivative $f' \in L^1([a,b])$ and*

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

Proof. Owing to Theorem 4.1.4, we assume without loss of generality that f is continuous. To show that f is differentiable a.e. in $[a, b]$, it is enough to show that

(a) $D^+f(x) < \infty$ for a.e. in $[a, b]$

(b) and $D^+f(x) \le D_-f(x)$ for a.e. in $[a, b]$.

Let $E_n := \{x \in [a, b] \mid D^+f(x) > n\}$ for each fixed $n \in \mathbb{N}$. Note that E_n is measurable and is a decreasing sequence of sets. Also,

$$\{x \in [a, b] \mid D^+f(x) = \infty\} = \cap_n E_n.$$

Now apply Lemma 4.1.3 to $g_n(x) := f(x) - nx$. Therefore, $E_n \subset \cup_k(a_k, b_k)$ and $f(b_k) - f(a_k) \ge n(b_k - a_k)$. Thus,

$$\mu^\star(E_n) \le \sum_k (b_k - a_k) \le \frac{1}{n}\sum_k [f(b_k) - f(a_k)] \le \frac{1}{n}[f(b) - f(a)].$$

Thus, $\mu^\star(E_n) \to 0$ as $n \to \infty$. Thus $\mu^\star(\{D^+f(x) = \infty\}) = 0$. Thus, $D^+f < \infty$ almost everywhere.

Now we shall prove that $D^+f(x) \le D_-f(x)$ for a.e. in $[a, b]$. It is enough to show this part because, by applying this result to $g(y) = -f(-y)$, we would get $D^-f(-y) \le D_+f(-y)$ for a.e. $-y \in [a, b]$. Hence,

$$D^+f(x) \le D_-f(x) \le D^-f(x) \le D_+f(x)$$

and all are equal since $D_+f(x) \le D^+f(x)$. Thus, it is sufficient to show that the set

$$E := \{x \in [a, b] \mid D^+f(x) > D_-f(x)\}$$

has outer measure zero. In fact, a similar argument will prove the result for every other combination of Dini derivatives. Let $p, q \in \mathbb{Q}$ such that $p > q$ and define

$$E_{p,q} := \{x \in [a, b] \mid D^+f(x) > p > q > D_-f(x)\}.$$

Note that $E = \cup_{\substack{p,q\in\mathbb{Q}\\ p>q}} E_{p,q}$. We will show that $\mu^\star(E_{p,q}) = 0$ which will imply that $\mu^\star(E) = 0$. To begin we assume a non-empty $E_{p,q}$ has $\mu^\star(E_{p,q}) \ne 0$, for a fixed $p, q \in \mathbb{Q}$ such that $p > q$, and arrive at a contradiction. We construct a Vitali cover of $E_{p,q}$. For any given $\varepsilon > 0$, by outer regularity, there is an open set $\Omega \supset E_{p,q}$ such that $\mu(\Omega) < \mu^\star(E_{p,q}) + \varepsilon$. For each $x \in E_{p.q}$, since Ω is open, there is an interval $[x-h, x] \subset \Omega$ such that

$$f(x) - f(x-h) < qh.$$

The collection of all such intervals, for each $x \in E_{p,q}$, forms a Vitali cover of $E_{p,q}$. Therefore, by Vitali covering lemma, we have finite disjoint subcollection $\{I_i\}_1^m$ from the Vitali cover such that

$$\mu^\star(E_{p,q} \setminus \cup_{i=1}^m I_i) < \varepsilon.$$

Therefore, we have

$$\sum_{i=1}^m (f(x_i) - f(x_i - h_i)) < q \sum_{i=1}^m h_i = q \sum_{i=1}^m \mu(I_i) < q\mu(\Omega) < q(\mu^\star(E_{p,q}) + \varepsilon).$$

Now let $A = E_{p,q} \cap (\cup_{i=1}^m \mathrm{Int}(I_i))$ and hence $E_{p,q} = A \cup (E_{p,q} \setminus \cup \mathrm{Int}(I_i))$. Thus, $\mu^\star(E_{p,q}) < \mu^\star(A) + \varepsilon$. We shall now construct a Vitali cover for A in terms of I_i. Note that each $y \in A$ is contained in $\mathrm{Int}(I_i)$ for some i. Choose $k > 0$ such that $[y, y+k] \subseteq I_i$ and

$$f(y+k) - f(y) > pk.$$

By Vitali covering lemma, there is finite disjoint collection of intervals $\{J_j\}_1^\ell$ each contained in I_i for some i such that

$$\mu^\star(A \setminus \cup_{j=1}^\ell J_j) < \varepsilon.$$

Set $B = A \cap (\cup_{j=1}^\ell J_j)$ and $A = B \cup (A \setminus \cup_{j=1}^\ell J_j)$. Hence, $\mu^\star(A) < \mu^\star(B) + \varepsilon$. Therefore, we have

$$\begin{aligned} \sum_{j=1}^\ell (f(y_j + k_j) - f(y_j)) &> p \sum_{j=1}^\ell k_j = p \sum_{j=1}^\ell \mu(J_j) = p\mu(\cup_j J_j) \\ &\geq p\mu^\star(B) > p(\mu^\star(A) - \varepsilon) > p(\mu^\star(E_{p,q}) - 2\varepsilon). \end{aligned}$$

Now, for each fixed i, we sum over all j such that $J_j \subset I_i$ to get the inequality

$$\sum_{J_j \subseteq I_i} (f(y_j + k_j) - f(y_j)) \leq f(x_i) - f(x_i - h_i),$$

due to the increasing nature of f and disjointness of J_j. This implies that $p(\mu^\star(E_{p,q}) - 2\varepsilon) < q(\mu^\star(E_{p,q}) + \varepsilon)$. Since ε is arbitrary, we have $p\mu^\star(E_{p,q}) \leq q\mu^\star(E_{p,q})$ which will contradict $q < p$ unless $\mu^\star(E_{p,q}) = 0$. Consequently, $\mu^\star(E) = 0$ and f is differentiable a.e. in $[a,b]$. Hence $f'(x)$ is defined a.e. in $[a,b]$. Set

$$g_k(x) = k\left(f(x + 1/k) - f(x)\right)$$

such that for all $x \geq b$, $f(x) = f(b)$[1]. Note that $g_k(x) \to f'(x)$ a.e. in $[a,b]$. Thus, f' is measurable, due to the measurability of g_k which follows from the measurability of f, a consequence of being a increasing function. Also, since f is increasing g_k are non-negative and hence f' is non-negative. Using Fatou's lemma, we have

$$\begin{aligned}
\int_a^b f' &\leq \liminf_{k\to\infty} \int_a^b g_k = \liminf_{k\to\infty} \left(k\int_a^b f(x + 1/k) - k\int_a^b f(x)\right) \\
&= \liminf_{k\to\infty} \left(k\int_{a+1/k}^{b+1/k} f(x) - k\int_a^b f(x)\right) \\
&= \liminf_{k\to\infty} \left(k\int_b^{b+1/k} f(x) - k\int_a^{a+1/k} f(x)\right) \quad (f \text{ constant for } x \geq b) \\
&= \liminf_{k\to\infty} \left(f(b) - k\int_a^{a+1/k} f(x)\right) \\
&\leq f(b) - f(a) \quad (f \text{ is increasing}).
\end{aligned}$$

□

Note that the above result also holds true for decreasing functions. Also, observe that for any two increasing functions their sum and difference are also differentiable a.e., but the difference is not necessarily increasing or decreasing. We wish to classify this class of functions which is the difference of two increasing functions.

4.2 Bounded Variation Functions

The problem of finding area under a graph lead to the notion of integration. An equally important problem is to find the length of curves. Let γ denote

[1]One could have chosen $f(x) = c$, for any $c \geq f(b)$, for all $x \geq b$, and obtain $c - f(a)$ but $f(b) - f(a)$ is the best bound one can obtain

a continuous curve in a metric space (X, d). Let the continuous function $\gamma : [a, b] \to X$ be the parametrisation of the curve γ with parametrised variable $t \in [a, b]$. Let P be the partition of the interval $[a, b]$, $a = t_0 \leq t_1 \leq \ldots \leq t_k = b$.

Definition 4.2.1. *The* length of a curve $\gamma : [a, b] \to X$ *on a metric space* (X, d) *is defined as*

$$L(\gamma) := \sup_P \left\{ \sum_{i=1}^{k} d(\gamma(t_i) - \gamma(t_{i-1})) \right\},$$

where the supremum is taken over all finite number of partitions P *of* $[a, b]$*. If* $L(\gamma) < +\infty$ *then the curve* γ *is said to be* rectifiable*.*

The length of the curve is defined as the supremum over the sum of length of all finite number of "line segments" approximating γ. If X is the usual Euclidean space with standard metric then the length of the curve has the form

$$L(\gamma) = \sup_P \left\{ \sum_{i=1}^{k} |[(\gamma(t_i))^2 - (\gamma(t_{i-1}))^2]^{1/2} \right\}.$$

A interesting questions one can ask at this juncture is: under what conditions on the function γ is the curve γ rectifiable? The length of a curve definition motivates the class of bounded variation functions.

Definition 4.2.2. *Let* $f : [a, b] \to \mathbb{R}(\mathbb{C})$ *be any real or complex valued function.*[2] *Let* P *be a partition of the interval* $[a, b]$, $a = x_0 \leq x_1 \leq \ldots \leq x_k = b$*. We define the* total variation *of* f *on* $[a, b]$*, denoted as* $V(f; [a, b])$*, as*

$$V(f; [a, b]) := \sup_P \left\{ \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \right\}.$$

We say f *is of bounded variation if* $V(f; [a, b]) < +\infty$ *and the class of all bounded variation function is denoted as* $BV([a, b])$*.*

Comparing the definition of bounded variation with curves in $\mathbb{C}$, we expect that any curve γ is rectifiable iff $\gamma : [a, b] \to \mathbb{C}$ is a function of bounded variation.

[2] not necessarily continuous as required for a curve γ

Example 4.2. Every constant function on $[a,b]$ belongs to $BV([a,b])$ and its total variation, $V(f;[a,b]) = 0$, is zero.

Lemma 4.2.3. *For any function f, $V(f;[a,b]) = 0$ iff f is a constant function on $[a,b]$*

Example 4.3. Any increasing function f on $[a,b]$ has the total variation $V(f;[a,b]) = f(b) - f(a)$. Consequently, if f is a bounded increasing function, then $f \in BV([a,b])$. For any partition $P = \{a = x_0 \leq \ldots \leq x_k = b\}$, we have

$$\begin{aligned}
\sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^{k} (f(x_i) - f(x_{i-1})) \\
&= f(x_1) - f(a) + f(x_2) - f(x_1) + \ldots + \\
&\quad + f(x_{k-1}) - f(x_{k-2}) + f(b) - f(x_{k-1}) \\
&= f(b) - f(a).
\end{aligned}$$

Thus, $V(f;[a,b]) = f(b) - f(a)$. Similarly, if f is decreasing on $[a,b]$ then $f \in BV([a,b])$ and $V(f;[a,b]) = f(a) - f(b)$.

The above example is very important and we will later see that every function of bounded variation can be decomposed in to increasing bounded functions, a result due to Jordan.

Example 4.4. The Cantor function f_C on $[0,1]$ is increasing and hence is in $BV([0,1])$. We already know f_C is uniformly continuous (cf. Appendix A).

Example 4.5. Any differentiable function $f : [a,b] \to \mathbb{R}$ such that f' is bounded (say by C) is in $BV([a,b])$. Using mean value theorem, we know that

$$f'(z) = \frac{|f(x) - f(y)|}{|x - y|} \quad \forall x, y \in [a,b] \text{ and } z \in [x,y].$$

Since the derivative is bounded, we get $|f(x) - f(y)| \leq C|x - y|$. Thus, for any partition $P = \{a = x_0 \leq \ldots \leq x_k = b\}$, we have

$$\sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| \leq C \sum_{i=1}^{k} |x_i - x_{i-1}| = C(b - a).$$

The above example is a particular case of the class of Lipschitz functions on $[a,b]$.

Definition 4.2.4. *A function $f : [a,b] \to \mathbb{R}(\mathbb{C})$ is said to be* Lipschitz *on $[a,b]$ if there exists a Lipschitz constant $C > 0$ such that*

$$|f(x) - f(y)| \leq C|x - y| \quad \forall x, y \in [a,b].$$

The space of all Lipschitz functions is denoted as $\mathrm{Lip}([a,b])$.

Exercise 85. Any Lipschitz function is uniformly continuous, $\mathrm{Lip}([a,b]) \subset C([a,b])$

Exercise 86. Every Lipschitz function is of bounded variation and

$$V(f;[a,b]) \leq C(b-a),$$

i.e., $\mathrm{Lip}([a,b]) \subset BV([a,b])$.

Exercise 87. Every element of $BV([a,b])$ is a bounded function on $[a,b]$.

Exercise 88. Show that the characteristic function $f = \chi_{\mathbb{Q}\cap[a,b]}$ on $[a,b]$ do not belong to $BV([a,b])$.

Proof. The idea behind the proof is that for each fixed $n \in \mathbb{N}$, we shall construct a partition $P = \{x_0, x_1, \ldots, x_n, x_{n+1}, x_{n+2}\}$ of $[a,b]$ such that

$$V(f;[a,b]) \geq \sum_{i=1}^{n+2} |f(x_i) - f(x_{i-1})| > n.$$

Choose $x_0 = a$. Let x_1 be an irrational between a and b. Choose x_2 to be an rational between x_1 and b. Proceeding this way till $x_{n+2} = b$, we will have a partition P whose successive points, excluding a and b, alternate between rational and irrational. Therefore,

$$\begin{aligned}
V(f;[a,b]) &\geq \sum_{i=1}^{n+2} |f(x_i) - f(x_{i-1})| \\
&\geq \sum_{i=2}^{n+1} |f(x_i) - f(x_{i-1})| \\
&= |f(x_2) - f(x_1)| + \ldots + |f(x_{n+1}) - f(x_n)| \\
&= |1 - 0| + |0 - 1| + \ldots = n.
\end{aligned}$$

Thus, $V(f;[a,b]) = \infty$. $\square$

We have already seen that the Cantor function f_C, which is (uniformly) continuous, is of bounded variation. But we do have functions which are continuous and not of bounded variation.

Exercise 89. Show that the following continuous function on $[0,1]$

$$f(x) = \begin{cases} x\sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

do not belong to $BV([0,1])$. More generally,

$$g(x) = \begin{cases} x^\alpha \sin(1/x^\beta) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is in $BV([0,1])$ iff $\alpha > \beta$.

Proof. For each $k \in \mathbb{N}$, note that

$$f\left(\frac{1}{k\pi}\right) = 0 \quad \text{and } f\left(\frac{1}{k\pi + \pi/2}\right) = \frac{(-1)^k}{k\pi + \pi/2}.$$

If we choose the points in our partition alternating between $1/k\pi$ and $\frac{1}{k\pi+\pi/2}$ then, for each $m \in \mathbb{N}$,

$$V(f;[a,b]) \geq \sum_{k=1}^{m} \left| 0 - \frac{(-1)^k}{k\pi + \pi/2} \right|.$$

Hence,

$$V(f;[a,b]) \geq \sum_{k=1}^{\infty} \frac{1}{k\pi + \pi/2} = \infty.$$

For each $k \in \mathbb{N}$, note that

$$g\left(\frac{1}{(k\pi)^{1/\beta}}\right) = 0 \quad \text{and } f\left(\frac{1}{(k\pi + \pi/2)^{1/\beta}}\right) = \frac{(-1)^k}{(k\pi + \pi/2)^{\alpha/\beta}}.$$

If we choose the points in our partition alternating between $1/(k\pi)^{1/\beta}$ and $\frac{1}{(k\pi+\pi/2)^{1/\beta}}$ then, for each $m \in \mathbb{N}$,

$$V(f;[a,b]) \geq \sum_{k=1}^{m} \left| 0 - \frac{(-1)^k}{(k\pi + \pi/2)^{\alpha/\beta}} \right|.$$

Hence,

$$V(f;[a,b]) \geq \sum_{k=1}^{\infty} \frac{1}{(k\pi+\pi/2)^{\alpha/\beta}}.$$

The series on the right converges iff $\alpha/\beta > 1$. Thus, $g \in BV([0,1])$ implies $\alpha > \beta$. The converse part needs a proof. □

Lemma 4.2.5. *Let $f : [a,b] \to \mathbb{R}$ be a given function. Let P denote the partition $P = \{a, x_1, \ldots, x_{n-1}, b\}$ of the interval $[a,b]$ and $P' = \{a, y_1, \ldots, y_{m-1}, b\}$ be a refinement of P, i.e., $P \subset P'$. Then*

$$\sum_{P} |f(x_i) - f(x_{i-1})| \leq \sum_{P'} |f(y_j) - f(y_{j-1})|.$$

Proof. We first prove the result by adding one point to the partition P and then invoke induction. Let $y \in P'$. If $y = x_i$, for some i, then the partition P remains unchanged. If $y \neq x_i$ for all i then $y \in (x_{k-1}, x_k)$ for some $k \in \{0, 1, \ldots, n\}$. Consider,

$$\begin{aligned}
\sum_{P} |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^{k-1} |f(x_i) - f(x_{i-1})| + |f(x_k) - f(x_{k-1})| \\
&\quad + \sum_{i=k+1}^{n} |f(x_i) - f(x_{i-1})| \\
&= \sum_{i=1}^{k-1} |f(x_i) - f(x_{i-1})| + \sum_{i=k+1}^{n} |f(x_i) - f(x_{i-1})| \\
&\quad + |f(x_k) - f(y) + f(y) - f(x_{k-1})| \\
&\leq \sum_{i=1}^{k-1} |f(x_i) - f(x_{i-1})| + |f(x_k) - f(y)| \\
&\quad + |f(y) - f(x_{k-1})| + \sum_{i=k+1}^{n} |f(x_i) - f(x_{i-1})| \\
&= \sum_{i=1}^{n+1} |f(x_i) - f(x_{i-1})| \quad \text{(by relabelling)}.
\end{aligned}$$

Similarly, adding each point of P' into the extended partition of P, we have our result. □

Exercise 90. Show that $BV([a,b])$ forms a vector space over $\mathbb{R}$. Also, if $f,g \in BV([a,b])$ then

(i) fg are in $BV([a,b])$.

(ii) $f/g \in BV([a,b])$ if $1/g$ is bounded on $[a,b]$.

Theorem 4.2.6. *Let $f : [a,b] \to \mathbb{R}$ be a function and let $c \in (a,b)$. If f belongs to both $BV([a,c])$ and $BV([c,b])$ then $f \in BV([a,b])$ and*

$$V(f;[a,b]) = V(f;[a,c]) + V(f;[c,b]).$$

Proof. Let P be any partition of $[a,b]$ and $P' = P \cup \{c\}$, relabelled in increasing order. P' being a refinement of P, arguing similar to the proof of Lemma 4.2.5, we get

$$\begin{aligned}\sum_P |f(x_i) - f(x_{i-1})| &\leq \sum_{P'} |f(x_i) - f(x_{i-1})| \\ &= \sum_{x_i \leq c} |f(x_i) - f(x_{i-1})| + \sum_{x_i \geq c} |f(x_i) - f(x_{i-1})| \\ &\leq V(f;[a,c]) + V(f;[c,b]).\end{aligned}$$

Hence, $V(f;[a,b]) \leq V(f;[a,c]) + V(f;[c,b])$. On the other hand, let P_1 and P_2 be a partition of $[a,c]$ and $[c,b]$, respectively. Then $P = P_1 \cup P_2$ gives a partition of $[a,b]$. Therefore,

$$\begin{aligned}\sum_{P_1} |f(x_i) - f(x_{i-1})| + \sum_{P_2} |f(x_i) - f(x_{i-1})| &= \sum_P |f(x_i) - f(x_{i-1})| \\ &\leq V(f;[a,b]).\end{aligned}$$

The above inequality is true for any arbitrary partition P_1 and P_2 of $[a,c]$ and $[c,b]$, respectively, Thus,

$$V(f;[a,c]) + V(f;[c,b]) \leq V(f;[a,b])$$

and we have equality as desired. □

Exercise 91. Show that if $f \in BV([a,b])$ then $f \in BV([c,d])$ for all subintervals $[c,d] \subset [a,b]$.

Let $f : \mathbb{R} \to \mathbb{R}(\mathbb{C})$ be any real or complex valued function. Let $BV_{\text{loc}}(\mathbb{R})$ denote the class of all $f : \mathbb{R} \to \mathbb{R}$ such that $f \in BV([a,b])$ for all $[a,b] \subset \mathbb{R}$.

Definition 4.2.7. *We define the* total variation *of f on $\mathbb{R}$, denoted as $V(f;\mathbb{R})$, as*

$$V(f;\mathbb{R}) := \sup_{[a,b]} V(f;[a,b]),$$

where the supremum is taken over all closed intervals $[a,b] \subset \mathbb{R}$. We say f is of bounded variation on $\mathbb{R}$ if $V(f;\mathbb{R}) < +\infty$ and denote the class as $BV(\mathbb{R})$.

Note that $BV(\mathbb{R}) \subset BV_{\text{loc}}(\mathbb{R})$ and the inclusion is strict.

Example 4.6. The function f

$$f(x) = \begin{cases} \frac{1}{1-x} & x \neq 1 \\ 0 & x = 1 \end{cases}$$

belongs to $BV(0,1)$ but do not belong to $BV([0,1])$.

Definition 4.2.8. *For $f \in BV([a,b])$, we define its* variation function*,*

$$V_f(x) = \begin{cases} V(f;[a,x]) & \forall x \in (a,b] \\ 0 & x = a \end{cases}$$

Lemma 4.2.9. *The variation function $V_f(x)$ corresponding to a function $f \in BV([a,b])$ is an increasing function.*

Proof. Let $x, y \in [a,b]$ be such that $x < y$. We claim that $V_f(x) \leq V_f(y)$. By, Theorem 4.2.6, we have

$$\begin{aligned} V(f;[a,y]) &= V(f;[a,x]) + V(f;[x,y]) \\ V(f;[a,y]) - V(f;[a,x]) &= V(f;[x,y]) \\ V_f(y) - V_f(x) &= V(f;[x,y]). \end{aligned}$$

Since $V(f;[x,y]) \geq 0$, we have $V_f(y) \geq V_f(x)$ and equality holds when f is constant on $[x,y]$. $\square$

Theorem 4.2.10 (Jordan Decomposition)**.** *Let $f : [a,b] \to \mathbb{R}$ be a real valued function. Then the following are equivalent:*

(i) $f \in BV([a,b])$

(ii) There exist two increasing functions $f_1, f_2 : [a,b] \to \mathbb{R}$ such that $f = f_1 - f_2$.

Proof. (ii) implying (i) is obvious, because any increasing function is in the vector space $BV([a,b])$. Conversely, let us prove (i) implies (ii). For a given $f \in BV([a,b])$ we know that V_f, the variation function, is increasing in $[a,b]$. Set $f_1 = V_f$ and $f_2 = V_f - f$. It only remains to show that f_2 is increasing. Let $x, y \in [a,b]$ be such that $x < y$. Consider,

$$\begin{aligned} f_2(y) - f_2(x) &= V_f(y) - f(y) - V_f(x) + f(x) \\ &= V_f(y) - V_f(x) - (f(y) - f(x)) \\ &= V(f;[x,y]) - (f(y) - f(x)) \\ &\geq V(f;[x,y]) - |f(y) - f(x)| \geq 0. \end{aligned}$$

Thus, f_2 is increasing and $f = f_1 - f_2$. □

Exercise 92. Show that in the above theorem one can, in fact, have strictly increasing functions f_1, f_2.

Proof. $f = g_1 - g_2$ where $g_1 := f_1 + x$ and $g_2 := f_2 + x$. □

Theorem 4.2.11 (Lebesgue Differentiation Theorem)**.** *If $f \in BV([a,b])$ then f is differentiable a.e. in $[a,b]$ and the derivative $f' \in L^1([a,b])$. Further,*

$$\int_a^b |f'| \leq V(f;[a,b]).$$

Proof. The fact that f is differentiable a.e. and $f' \in L^1[a,b]$ follows from the Jordan decomposition (Theorem 4.2.10) and Theorem 4.1.6. Also, by Lemma 4.2.9, V_f is an increasing function. Thus, again by Theorem 4.1.6, V_f is differentiable a.e. and

$$\int_a^b V_f'(x)\,dx \leq V_f(b) - V_f(a) = V_f(b) = V(f;[a,b]).$$

For any $x, y \in [a,b]$, we have

$$\begin{aligned} V_f(y) - V_f(x) &= V(f;[x,y]) \\ &\geq |f(y) - f(x)| \geq f(y) - f(x). \end{aligned}$$

Thus, $f' \leq V_f'$ and $|f'| \leq |V_f'| = V_f'$. Therefore,

$$\int_a^b |f'(x)|\,dx \leq \int_a^b V_f'(x)\,dx \leq V(f;[a,b]).$$

□

Exercise 93. Give an example of a function $f : [0,1] \to \mathbb{R}$ such that $f \notin BV([0,1])$ but f is differentiable everywhere in $[0,1]$.

Proof. Let

$$f(x) = \begin{cases} x^2 \sin(1/x^2) & x \in (0,1] \\ 0 & x = 0 \end{cases}$$

then

$$f'(x) = \begin{cases} 2x \sin(1/x^2) - \frac{2}{x}\cos(1/x^2) & x \in (0,1] \\ 0 & x = 0. \end{cases}$$

□

The derivative f' considered in above exercise is not in $L^1([0,1])$. The L^1 belonging is very crucial to prove the converse of Lebesgue differentiation theorem. A kind of converse of Lebesgue differentiation theorem is proved in Theorem 4.4.5.

4.3 Derivative of an Integral

We now have enough tools to answer the first question posed in the beginning of this chapter. Let $f \in L^1([a,b])$ and set

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

We have already seen in Exercise 68 that F is continuous. We now show that F has bounded variation. In fact, a stronger statement is true, F is absolutely continuous, which we will prove once we introduce the definition of absolute continuity of a function.

Lemma 4.3.1. *If $f \in L^1([a,b])$ then the continuous function $F \in BV([a,b])$.*

Proof. Consider

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

Hence $V(f;[a.b]) < \infty$. □

Lemma 4.3.2. *If $f \in L^1([a,b])$ and $F \equiv 0$ on $[a,b]$ then $f = 0$ a.e. on $[a,b]$.*

Proof. We shall prove by contradiction. Let $E := \{x \in [a,b] \mid f(x) > 0\}$. Assume E is of non-zero measure. By inner regularity there exists a closed set $\Gamma \subset E$ such that $\mu(E \setminus \Gamma) < \varepsilon$ for any given $\varepsilon > 0$. Hence $\mu(\Gamma) > 0$. Set $\Omega = (a,b) \setminus \Gamma$. Since F is identically zero, we have

$$0 = F(b) = \int_a^b f(t)\,dt = \int_\Gamma f + \int_\Omega f.$$

Thus,

$$\int_\Omega f = -\int_\Gamma f \neq 0.$$

Since Ω is open, $\Omega = \cup_{i=1}^\infty (a_i, b_i)$ is a disjoint union of open intervals. Then,

$$0 \neq \int_\Omega f = \sum_{i=1}^\infty \int_{a_i}^{b_i} f.$$

Therefore, for some $k \in \mathbb{N}$, we have

$$0 \neq \int_{a_k}^{b_k} f = \int_a^{b_k} f - \int_a^{a_k} f = F(b_k) - F(a_k).$$

Thus, either $F(b_k) \neq 0$ or $F(a_k) \neq 0$ which contradicts the hypothesis on F. Similar argument is valid for the set E on which $f < 0$. Hence proved. $\square$

Theorem 4.3.3. *Let $f \in L^1([a,b])$ and $c \in \mathbb{R}$. Set*

$$F(x) = c + \int_a^x f(t)\,dt.$$

Then $c = F(a)$ and $F' = f$ a.e. on $[a,b]$.

Proof. The fact that $c = F(a)$ is obvious. By Lemma 4.3.1, we have $F \in BV([a,b])$. By Lebesgue differentiation theorem, F is differentiable a.e. It is required to show that $F' = f$ a.e. We prove by cases. First let us assume f is bounded, i.e., $\|f\|_\infty < \infty$. Extend F as $F(x) = F(b)$, for all $x \geq b$. Set

$$g_k(x) := k\left(F(x + 1/k) - F(x)\right) = k \int_x^{x+1/k} f(t)\,dt.$$

Note that $g_k(x) \to F'(x)$ a.e. in $[a, b]$. Since f is bounded, g_k's are all uniformly bounded and supported inside $[a, b]$. Using BCT, for any $d \in [a, b]$, we have

$$\begin{aligned}
\int_a^d F' &= \lim_{k\to\infty} \int_a^d g_k = \lim_{k\to\infty} \left(k \int_a^d F(x + 1/k) - k \int_a^d F(x) \right) \\
&= \lim_{k\to\infty} \left(k \int_{a+1/k}^{d+1/k} F(x) - k \int_a^d F(x) \right) \\
&= \lim_{k\to\infty} \left(k \int_d^{d+1/k} F(x) - k \int_a^{a+1/k} F(x) \right).
\end{aligned}$$

Consider, for any $e \in [a, b]$,

$$\begin{aligned}
\left| F(e) - \lim_{h\to 0} \frac{1}{h} \int_e^{e+h} F(x)\,dx \right| &\leq \lim_{h\to 0} \frac{1}{h} \int_e^{e+h} |F(e) - F(x)|\,dx \\
&\leq \sup_{x\in[e,e+h]} |F(e) - F(x)|.
\end{aligned}$$

By continuity of F (cf. Exercise 68), $F(e + h) \to F(e)$ as $h \to 0$. Hence, we have

$$\lim_{h\to 0} \frac{1}{h} \int_e^{e+h} F(x)\,dx = F(e).$$

Therefore,

$$\int_a^d F' = F(d) - F(a) = \int_a^d f(t)\,dt.$$

Thus, for all $d \in [a, b]$, we have

$$\int_a^d (F' - f) = 0$$

and by Lemma 4.3.2, we have $F' - f = 0$ a.e. on $[a, b]$, i.e., $F' = f$ a.e. on $[a, b]$.

It now remains to prove the result for a unbounded function. Without loss of generality, we assume f is non-negative. Then F is increasing on $[a, b]$ and, by Theorem 4.1.6, we have

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

Let f_k be the truncation of f at k level, i.e.,

$$f_k(x) = \begin{cases} f(x) & f(x) \leq k \\ k & f(x) > k. \end{cases}$$

Each f_k is a bounded by k and they converge a.e. to f. Also, $f - f_k$ is a non-negative function. Define

$$G_k(x) = \int_a^x (f(t) - f_k(t))\, dt.$$

Note that G_k is increasing function on $[a, b]$ and hence is in $BV([a, b])$. Thus, G_k is differentiable a.e. Since $\{f_k\}$ are each bounded, we have $f_k(x) = F_k'(x)$ a.e. where

$$F_k(x) = c + \int_a^x f_k(t)\, dt.$$

Therefore,

$$G_k'(x) = F'(x) - F_k'(x) = F'(x) - f_k(x).$$

Since G_k's are increasing its derivative is non-negative and hence, we have $F'(x) \geq f_k(x)$ a.e. Consequently, $F'(x) \geq f(x)$ a.e. Thus,

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

and we have equality above, since other inequality holds as noted above. Therefore,

$$\int_a^b (F'(x) - f(x))\, dx = 0$$

and for $F' - f \geq 0$, integral zero implies that $F'(x) = f(x)$ a.e. on $[a, b]$. $\square$

Recall the definition of derivative of F,

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt.$$

Thus, we may reformulate the demand $F'(x) = f(x)$ a.e. on $[a, b]$ as

$$\lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)\, dt = f(x) \quad \text{for a.e. } x \in [a, b].$$

Note that the integral (along with the fraction) on the LHS is the "average" or "mean" of f over $[x, x+h]$ and the equality says that the limit of averages of f around a interval I of x converges to the value of f at x, as the measure of interval I tends to zero. The theorem proved above validates this result for all $f \in L^1$ in the one dimension case. This reformulation, in terms of averages, helps in stating the problem in higher dimensions. Thus, in higher dimension, we ask the question: For all $f \in L^1(\mathbb{R}^n)$, do we have

$$\lim_{\substack{\mu(B)\to 0 \\ x\in B}} \frac{1}{\mu(B)} \int_B f(t)\,dt = f(x) \quad \text{for a.e. } x \in \mathbb{R}^n \tag{4.3.1}$$

where B is any ball in $\mathbb{R}^n$ containing x? Note that if f is continuous at x then (4.3.1) holds true because for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(t)| < \varepsilon$ whenever $|x - t| < \delta$ and, hence, for any ball B with radius less than $\delta/2$ that contains x, we have

$$\left| f(x) - \frac{1}{\mu(B)} \int_B f(t)\,dt \right| \leq \frac{1}{\mu(B)} \int_B |f(x) - f(t)|\,dt < \varepsilon.$$

Theorem 4.3.4 (Lebesgue-Besicovitch Differentiation Theorem)**.** *Let μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^1_{loc}(\mathbb{R}^n, \mu)$. Then*

$$\lim_{h\to 0} \frac{1}{\mu(B_h(x))} \int_{B_h(x)} f\,d\mu = f(x)$$

for μ a.e. $x \in \mathbb{R}^n$.

For the case when μ is a Lebesgue measure, we define the *precise representative* of f as

$$f^\star(x) := \begin{cases} \lim_{h\to 0} \frac{1}{\mu(B)} \int_B f\,dx & \text{if the limit exists} \\ 0 & \text{otherwise} \end{cases}$$

By Lebesgue-Besicovitch differentiation theorem, $f = f^\star$ a.e., i.e., they are in the same equivalence class. If $f = g$ a.e then

$$\lim_{h\to 0} \frac{1}{\mu(B)} \int_B f(t)\,dt = \lim_{h\to 0} \frac{1}{\mu(B)} \int_B g(t)\,dt$$

whenever the limit exists. Thus, $f^\star = g^\star$ for all $x \in \mathbb{R}^n$.

Definition 4.3.5. *Let $1 \leq p < \infty$, μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^p_{loc}(\mathbb{R}^n, \mu)$. A point $x \in \mathbb{R}^n$ is said to be a* Lebesgue point *of f w.r.t μ if*

$$\lim_{h \to 0} \frac{1}{\mu(B_h(x))} \int_{B_h(x)} |f - f(x)|^p \, d\mu = 0.$$

The set of all Lebesgue points of f w.r.t μ is called the Lebesgue set *of f w.r.t μ.*

In this terminology, Lebesgue differentiation theorem says that the complement of Lebesgue set is of measure zero. Further, the set of all continuity points of f is contained in the Lebesgue set.

Corollary 4.3.6. *Let $1 \leq p < \infty$, μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^p_{loc}(\mathbb{R}^n, \mu)$. The complement of the Lebesgue set of f w.r.t μ has measure zero w.r.t μ.*

Corollary 4.3.7. *If μ is the Lebesgue measure then result with balls having centre at x is also true for any ball containing x, i.e.,*

$$\lim_{\substack{\mu(B) \to 0 \\ x \in B}} \frac{1}{\mu(B)} \int_B |f(t) - f(x)|^p \, dt = 0 \quad \text{for a.e. } x \in \mathbb{R}^n.$$

4.4 Absolute Continuity and FTC

In this section we wish to identify the class of functions f for which

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

This is the second question we hoped to answer in the beginning of the chapter. Note that if $f \in BV([a,b])$ then, by Lebesgue differentiation theorem, f is differentiable and the derivative is Lebesgue integral. So the question reducing to asking: For any $f \in BV([a,b])$, do we have

$$\int_a^b f'(x) \, dx = f(b) - f(a)?$$

Unfortunately, the answer is a “No”, as seen in the example below.

Example 4.7. Recall that the Cantor function $f_C \in BV([0,1])$, since f_C is increasing. Outside of the Cantor set C, f_C is constant and hence $f_C' = 0$ a.e. on $[0,1]$. Therefore,

$$\int_0^1 f_C' = 0$$

but $f_C(1) - f_C(0) = 1 - 0 = 1$.

This motivates us to look for a sub-class of bounded variation functions for which *fundamental theorem of calculus* (FTC) is true.

Definition 4.4.1. *A function $f : [a,b] \to \mathbb{R}$ is said to be* absolutely continuous *on $[a,b]$ if for every $\varepsilon > 0$ there exist a $\delta > 0$ such that*

$$\sum_i |f(y_i) - f(x_i)| < \varepsilon.$$

for any disjoint collection (finite or countable) of subintervals $\{(x_i, y_i)\}$ of $[a,b]$ with

$$\sum_i |y_i - x_i| < \delta.$$

Let $AC([a,b])$ denote the set of all absolutely continuous functions on $[a,b]$.

Exercise 94. Show that $AC([a,b])$ forms a vector space over $\mathbb{R}$ or $\mathbb{C}$. Also, show that $f, g \in AC([a,b])$ then $fg \in AC([a,b])$.

Exercise 95. If $f \in AC([a,b])$ then $|f| \in AC([a,b])$.

Exercise 96. Show that $AC([a,b]) \subset C([a,b])$. The inclusion is proper. Show that the Cantor function, which is continuous, $f_C \notin AC([a.b])$.

Example 4.8. Every constant function on $[a,b]$ is in $AC([a,b])$.

Example 4.9. For any $f \in L^1([a,b])$, $F(x) = \int_a^x f(t)\,dt$ is absolutely continuous in $[a,b]$. Let $\varepsilon > 0$ be given. By Proposition 3.4.6, we have $\delta > 0$ such that

$$\int_E |f| < \varepsilon \text{ whenever } \mu(E) < \delta.$$

Let us pick a collection of disjoint subintervals $\{(x_i, y_i)\} \subset [a,b]$ such that $\sum_i |y_i - x_i| < \delta$. Consider

$$\sum_i |F(y_i) - F(x_i)| \leq \sum_i \int_{x_i}^{y_i} |f(t)|\,dt = \int_{\cup_i (x_i, y_i)} |f(t)|\,dt < \varepsilon.$$

Example 4.10. $\text{Lip}([a,b]) \subset AC([a,b])$. This inclusion is proper.

Lemma 4.4.2. *Every absolutely continuous function is of bounded variation, i.e., $AC([a,b]) \subset BV([a,b])$. Consequently, if $f \in AC([a,b])$ then f is differentiable a.e. in $[a,b]$ and $f \in L^1([a,b])$.*

Proof. Let $f \in AC([a,b])$ and $\delta > 0$ be such that for any disjoint collection (finite or countable) of subintervals $\{(x_i, y_i)\}$ of $[a,b]$ with

$$\sum_i |y_i - x_i| < \delta$$

we have

$$\sum_i |f(y_i) - f(x_i)| < 1.$$

Let M denote the smallest integer such that $(b-a)/\delta \leq M$. Let P be any partition of $[a,b]$. We refine the partition P into P' such that P' has precisely M intervals and hence each of the interval has length less than δ. Therefore, for each subinterval of the partition P', we have $|f(x_i) - f(x_{i-1})| < 1$. Thus,

$$\sum_{P'} |f(x_i) - f(x_{i-1})| < M.$$

But by Lemma 4.2.5, we have

$$\sum_{P} |f(x_i) - f(x_{i-1})| \leq \sum_{P'} |f(x_i) - f(x_{i-1})| < M.$$

Hence, $V(f;[a,b]) < M$ and $f \in BV([a,b])$. $\square$

The inclusion proved above is proper. The Cantor function f_C on $[0,1]$ is in $BV([0,1]$ and $f_C \notin AC([0,1])$. We now define a class of functions which are complementary in nature to absolutely continuous functions.

Definition 4.4.3. *A function $f : [a,b] \to \mathbb{R}$ is said to be* singular *if f is differentiable a.e. in $[a,b]$ and $f' = 0$ a.e. in $[a,b]$.*

Example 4.11. The Cantor function f_C is singular on $[0,1]$.

In fact, any non-constant function is either singular or absolutely continuous.

Theorem 4.4.4. *If $f : [a, b] \to \mathbb{R}$ is both absolutely continuous and singular, then f is constant.*

Proof. It is enough to show that $f(a) = f(c)$ for all $c \in (a, b]$. Fix a $c \in (a, b]$. Due to the singular nature of f, we have a measurable set $E \subset (a, c)$ such that $\mu(E) = c - a$ and $f'(x) = 0$ on E. Due to the absolute continuity of f, for every $\varepsilon > 0$ there exist a $\delta > 0$ such that

$$\sum_i |f(y_i) - f(x_i)| < \frac{\varepsilon}{2}.$$

for any disjoint collection (finite or countable) of subintervals $\{(x_i, y_i)\}$ of $[a, b]$ with

$$\sum_i |y_i - x_i| < \delta.$$

We now construct a Vitali cover for E. For each $x \in E$, f is differentiable and derivative is zero. We choose the interval $[x, x + h] \subset [a, c]$ such that

$$|f(x + h) - f(x)| < \frac{\varepsilon}{2(c - a)} h.$$

By Vitali covering lemma, we can find a finite disjoint collection of closed intervals $\{I_i = [x_i, x_i + h_i]\}_1^k$ such that

$$\mu^\star(E \setminus \cup_{i=1}^k I_i) < \delta.$$

Note that $[a, c]$ is an interval and $\{x_i, x_i + h_i\}$. for all $i = 1, 2, \dots, k$ is a disjoint collection of intervals in $[a, c]$. Thus, relabelling as $\{x_0 = a, x_1, x_1 + h_1, x_2, x_2 + h_2, \dots, x_k, x_k + h_k, c = x_{k+1}\}$ and setting $h_0 = 0$, we have that

$$\sum_{i=0}^k |x_{i+1} - (x_i + h_i)| < \delta.$$

Hence, by absolute continuity, we have

$$\sum_{i=0}^k |f(x_{i+1}) - f(x_i + h_i)| < \frac{\varepsilon}{2}$$

Also,

$$\sum_{i=1}^k |f(x_i + h_i) - f(x_i)| < \frac{\varepsilon}{2(c - a)} \sum_{i=1}^k h_i < \frac{\varepsilon}{2(c - a)}(c - a) = \frac{\varepsilon}{2}.$$

Now, consider

$$\begin{aligned}
|f(c) - f(a)| &= \left| \sum_{i=0}^{k} f(x_{i+1}) - f(x_i + h_i) + \sum_{i=1}^{k} f(x_i + h_i) - f(x_i) \right| \\
&\leq \sum_{i=0}^{k} |f(x_{i+1}) - f(x_i + h_i)| + \sum_{i=1}^{k} |f(x_i + h_i) - f(x_i)| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Since the choice of ε is arbitrary, we have $|f(c) - f(a)| = 0$ and $f(c) = f(a)$. $\square$

Theorem 4.4.5. *If $f : [a, b] \to \mathbb{R}$ then the following are equivalent:*

(i) $f \in AC([a, b])$

(ii) f is differentiable a.e.,$f' \in L^1([a, b])$ and

$$f(x) - f(a) = \int_a^x f'(t)\, dt \quad x \in [a, b].$$

Proof. We first prove (ii) implies (i). Let $\varepsilon > 0$ be given. By Proposition 3.4.6, since $f' \in L^1([a, b])$, there exists a $\delta > 0$ such that

$$\int_E |f'| < \varepsilon \quad \text{whenever } \mu(E) < \delta.$$

Consider any disjoint collection of subintervals of $[x_i, y_i] \subset [a, b]$ such that $\sum_i |y_i - x_i| < \delta$. We claim that $\sum_i |f(y_i) - f(x_i)| < \varepsilon$. Consider

$$\sum_i |f(y_i) - f(x_i)| = \sum_i \left| \int_{x_i}^{y_i} f'(t)\, dt \right| \leq \sum_i \int_{x_i}^{y_i} |f'| < \varepsilon.$$

Conversely, let $f \in AC([a, b])$ then $f \in BV([a, b])$. Thus, by Lebesgue differentiation theorem, f is differentiable a.e. and $f' \in L^1([a, b])$. Define

$$F(x) := \int_a^x f'(t)\, dt$$

By Example 4.9, $F \in AC([a,b])$ and hence $g = f - F \in AC([a,b])$. By Theorem 4.3.3, we have $F' = f'$ a.e. on $[a,b]$. Thus, $g' = 0$ a.e., hence g is singular. Therefore, g is constant, $g \equiv c$ and

$$f(x) = c + F(x) = c + \int_a^x f'(t)\,dt.$$

Thus, $c = f(a)$ and

$$f(x) - f(a) = \int_a^x f'(t)\,dt.$$

□

The implication (ii) implies (i) is a kind of converse to Lebesgue differentiation theorem (with additional hypothesis). In fact, the exact converse of Lebesgue differentiation theorem, i.e., f' exists and $f' \in L^1$ implies that $f \in AC \subset BV$, is true, but requires more observation, viz. Banach-Zaretsky theorem.

Corollary 4.4.6. *If $f \in BV([a,b])$ then f can be decomposed as $f = f_a + f_s$ where $f_a \in AC([a,b])$ and f_s is singular on $[a,b]$). Moreover, f_a and f_s are unique up to additive constants and*

$$\int_a^x f'(t)\,dt = f_a(x).$$

Proof. Since $f \in BV([a,b])$, by Lebesgue differentiation theorem, f' exists and $f' \in L^1([a,b])$. Define

$$f_a := \int_a^b f'(x)\,dx.$$

By (b), $f_a \in AC([a,b])$. By, the derivative of an integral we have, $f_a' = f'$ a.e. on $[a,b]$. Define $f_s := f - f_a$. Thus, $f_s' = 0$ a.e. and hence f_s is singular.

Let $f = g+h$ where $g \in AC([a,b])$ and h is singular. Then, $g+h = f_a+f_s$. Hence $g - f_a = f_s - h$. LHS is in $AC([a,b])$ and RHS is singular. Therefore they must be equal to some constant c. $g = f_a + c$ and $h = f_s - c$.

□

Exercise 97 (Integration by parts). Show that if $f, g \in AC([a,b])$ then

$$\int_a^b f(x)g'(x)\,dx + \int_a^b f'(x)g(x)\,dx = f(b)g(b) - f(a)g(a).$$

Theorem 4.4.7. *If $f \in AC([a,b])$ then*

$$V(f;[a,b]) = \int_a^b |f'|.$$

Proof. By Lemma 4.4.2, $f \in BV[a,b]$. Thus, by Theorem 4.2.11, $\int |f'| \leq V(f;[a,b])$. Consider any partition $P := \{a, x_1, \ldots, x_n, b\}$ of $[a,b]$. Since $f \in AC[x_{i-1}, x_i]$ for all $i = 1, \ldots, n+1$, we have from Theorem 4.4.5 that

$$f(x_i) - f(x_{i-1}) = \int_{x_{i-1}}^{x_i} f'.$$

Thus,

$$\sum_P |f(x_i) - f(x_{i-1})| \leq \sum_P \int_{x_{i-1}}^{x_i} |f'| = \int_a^b |f'|$$

and taking supremum over all partitions we establish $V(f;[a,b]) \leq \int_a^b |f'|$. Thus, equality holds. □

Theorem 4.4.8. *Let $f \in AC([a,b])$. Then $\mu(f(E)) = 0$ for all $E \subseteq [a,b]$ such that $\mu(E) = 0$.*

Proof. Let $E \subseteq (a,b)$ be such that $\mu(E) = 0$. Note that we are excluding the end-points because $\{f(a), f(b)\}$ is measure zero subset of $f(E)$. Since $f \in AC([a,b])$, for every given $\varepsilon > 0$, there exists a $\delta > 0$ such that for every sub-collection of disjoint intervals $\{(x_i, y_i)\} \subset [a,b]$ with $\sum_i (y_i - x_i) < \delta$, we have $\sum_i |f(y_i) - f(x_i)| < \varepsilon$. By outer regularity of E, there is an open set $\Omega \supset E$ such that $\mu(\Omega) < \delta$. Wlog, we may assume $\Omega \subset (a,b)$, because otherwise we consider the intersection of Ω with (a,b). But $\Omega = \cup_i (x_i, y_i)$, a disjoint countable union of open intervals and

$$\sum_i |y_i - x_i| = \mu(\Omega) < \delta.$$

Consider,

$$\begin{aligned} \mu^\star(f(E)) &\leq \mu^\star(f(\Omega)) = \mu^\star(f(\cup_i (x_i, y_i))) \\ &= \mu^\star(\cup_i f(x_i, y_i)) \leq \sum_i \mu^\star(f(x_i, y_i)). \end{aligned}$$

Let $c_i, d_i \in (x_i, y_i)$ be points such that $f(c_i)$ and $f(d_i)$ is the minimum and maximum, respectively, of f on (x_i, y_i). Note that $|d_i - c_i| \leq |y_i - x_i|$ and hence $\sum_i |d_i - c_i| < \delta$. Then,

$$\sum_i \mu^\star(f((x_i, y_i))) = \sum_i |f(d_i) - f(c_i)| < \varepsilon.$$

Thus, $\mu^\star(f(E)) = 0$ and $f(E)$ is measurable set. $\square$

Appendices

Appendix A

Cantor Set and Cantor Function

Let us construct the Cantor set which plays a special role in analysis.

Consider $C_0 = [0, 1]$ and trisect C_0 and remove the middle open interval to get C_1. Thus, $C_1 = [0, 1/3] \cup [2/3, 1]$. Repeat the procedure for each interval in C_1, we get

$$C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

Repeating this procedure at each stage, we get a sequence of subsets $C_i \subseteq [0, 1]$, for $i = 0, 1, 2, \ldots$. Note that each C_k is a compact subset, since it is a finite union of compact sets. Moreover,

$$C_0 \supset C_1 \supset C_2 \supset \ldots \supset C_i \supset C_{i+1} \supset \ldots.$$

The Cantor set C is the intersection of all the nested C_i's, $C = \cap_{i=0}^{\infty} C_i$.

Lemma A.0.1. *C is compact.*

Proof. C is countable intersection of closed sets and hence is closed. $C \subset [0, 1]$ and hence is bounded. Thus, C is compact. □

The Cantor set C is non-empty, because the end-points of the closed intervals in C_i, for each $i = 0, 1, 2, \ldots$, belong to C. In fact, the Cantor set cannot contain any interval of positive length.

Lemma A.0.2. *For any $x, y \in C$, there is a $z \notin C$ such that $x < z < y$. (Disconnected)*

Proof. If $x, y \in C$ are such that $z \in C$ for all $z \in (x, y)$, then we have the open interval $(x, y) \subset C$. It is always possible to find i, j such that

$$\left(\frac{j}{3^i}, \frac{j+1}{3^i}\right) \subseteq (x, y)$$

but does not belong $C_i \supset C$. □

We show in example 2.5, that C has length zero. Since C is non-empty, how 'big' is C? The number of end-points sitting in C are countable. But C has points other than the end-points of the closed intervals C_i for all i. For instance, $1/4$ (not an end-point) will never belong to the the intervals being removed at every step i, hence is in C. There are more! $3/4$ and $1/13$ are all in C which are not end-points of removed intervals. It is easy to observe these by considering the ternary expansion characterisation of C. Consider the ternary expansion of every $x \in [0, 1]$,

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} = 0.a_1a_2a_3\ldots_3 \quad \text{where } a_i = 0, 1 \text{ or } 2.$$

The decomposition of x in ternary form is not unique[1]. For instance, $1/3 = 0.1_3 = 0.022222\ldots_3$, $2/3 = 0.2_3 = 0.1222\ldots_3$ and $1 = 0.222\ldots_3$. At the C_1 stage, while removing the open interval $(1/3, 2/3)$, we are removing all numbers whose first digit in ternary expansion (in all possible representations) is 1. Thus, C_1 has all those numbers in $[0, 1]$ whose first digit in ternary expansion is not 1. Carrying forward this argument, we see that for each i, C_i contains all those numbers in $[0, 1]$ with digits upto ith place, in ternary expansion, is not 1. Thus, for any $x \in C$,

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} = 0.a_1a_2a_3\ldots_3 \quad \text{where } a_i = 0, 2.$$

Lemma A.0.3. *C is uncountable.*

Proof. Use Cantor's diagonal argument to show that the set of all sequences containing 0 and 2 is uncountable. □

[1] This is true for any positional system. For instance, $1 = 0.99999\ldots$ in decimal system

Cantor Function

We shall now define the *Cantor function* $f_C : C \to [0,1]$ as,

$$f_C(x) = f_C\left(\sum_{i=1}^{\infty} \frac{a_i}{3^i}\right) = \sum_{i=1}^{\infty} \frac{a_i}{2} 2^{-i}.$$

Since $a_i = 0$ or 2, the function replaces all 2 occurrences with 1 in the ternary expansion and we interpret the resulting number in binary system. Note, however, that the Cantor function f_C is not injective. For instance, one of the representation of $1/3$ is $0.0222\ldots_3$ and $2/3$ is 0.2. Under f_C they are mapped to $0.0111\ldots_2$ and 0.1_2, respectively, which are different representations of the same point. Since f_C is same on the end-points of the removed interval, we can extend f_C to $[0,1]$ by making it constant along the removed intervals.

Alternately, one can construct the Cantor function step-by-step as we remove middle open intervals to get C_i. Consider f_1 to be a function which takes the constant value $1/2$ in the removed interval $(1/3, 2/3)$ and is linear on the remaining two intervals such that f_1 is continuous. In the second stage, the function f_2 coincides with f_1 in $1/3, 2/3$, takes the constant value $1/4$ and $3/4$ on the two removed intervals and is linear in the remaining four intervals such that f_2 is continuous. Proceeding this way we have a sequence of monotonically increasing continuous functions $f_k : [0,1] \to [0,1]$. Moreover, $|f_{k+1}(x) - f_k(x)| < 2^{-k}$ for all $x \in [0,1]$ and f_k converges uniformly to $f_C : [0,1] \to [0,1]$.

Exercise 98. The Cantor function $f_C : [0,1] \to [0,1]$ is uniformly continuous, monotonically increasing and is differentiable a.e. and $f_C' = 0$ a.e.

Exercise 99. The function f_C is not absolutely continuous.

Generalised Cantor Set

We generalise the idea behind the construction of Cantor sets to build Cantor-like subsets of $[0,1]$. Choose a sequence $\{a_k\}$ such that $a_k \in (0, 1/2)$ for all k. In the first step we remove the open interval $(a_1, 1-a_1)$ from $[0,1]$ to get C_1. Hence $C_1 = [0, a_1] \cup [1-a_1, 1]$. Let

$$C_1^1 := [0, a_1] \quad \text{and } C_1^2 := [1-a_1, 1].$$

Hence, $C_1 = C_1^1 \cup C_1^2$. Note that C_1^i are sets of length a_1 carved out from the end-points of C_0. We repeat step one for each of the end-points of C_1^i of length $a_1 a_2$. Therefore, we get four sets

$$C_2^1 := [0, a_1 a_2] \quad C_2^2 := [a_1 - a_1 a_2, a_1],$$
$$C_2^3 := [1 - a_1, 1 - a_1 + a_1 a_2] \quad C_2^4 := [1 - a_1 a_2, 1].$$

Define $C_2 = \cup_{i=1}^4 C_2^i$. Each C_2^i is of length $a_1 a_2$. Note that $a_1 a_2 < a_1$. Repeating the procedure successively for each term in the sequence $\{a_k\}$, we get a sequence of sets $C_k \subset [0,1]$ whose length is $2^k a_1 a_2 \dots a_k$. The "generalised" Cantor set C is the intersection of all the nested C_k's, $C = \cap_{k=0}^\infty C_k$ and each $C_k = \cup_{i=1}^{2^k} C_k^i$. Note that by choosing the constant sequence $a_k = 1/3$ for all k gives the Cantor set defined in the beginning of this Appendix. Similar arguments show that the generalised Cantor set C is compact. Moreover, C is non-empty, because the end-points of the closed intervals in C_k, for each $k = 0, 1, 2, \dots$, belong to C.

Lemma A.0.4. *For any $x, y \in C$, there is a $z \notin C$ such that $x < z < y$.*

Lemma A.0.5. *C is uncountable.*

We show in example 2.12, that C has length $2^k a_1 a_2 \dots a_k$.

The interesting fact about generalised Cantor set is that it can have non-zero "length".

Proposition A.0.6. *For each $\alpha \in [0, 1)$ there is a sequence $\{a_k\} \subset (0, 1/2)$ such that*

$$\lim_k 2^k a_1 a_2 \dots a_k = \alpha.$$

Proof. Choose $a_1 \in (0, 1/2)$ such that $0 < 2a_1 - \alpha < 1$. Use similar arguments to choose $a_k \in (0, 1/2)$ such that $0 < 2^k a_1 a_2 \dots a_k - \alpha < 1/k$. $\square$

Generalised Cantor Function

We shall define the generalised Cantor function f_C on the generalised Cantor set C. Define the function $f_0 : [0,1] \to [0,1]$ as $f_0(x) = x$. f_0 is continuous on $[0,1]$. We define $f_1 : [0,1] \to [0,1]$ such that f is linear and continuous on C_1^i, and $1/2$ on $[a_1, 1 - a_1]$, the closure of removed open interval at first stage. We define $f_k : [0,1] \to [0,1]$ continuous $f_k(0) = 0, f_k(1) = 1$ such that $f_k(x) = i/2^k$ on the removed interval immediate right to C_k^i.

Theorem A.0.7. *Each f_k is continuous, monotonically non-decreasing and uniformly converges to some $f_C : [0, 1] \to [0, 1]$.*

Thus, f_C being uniform limit of continuous function is continuous and is the called the generalised Cantor function.

Bibliography

[Pug04] Charles Chapman Pugh, *Real mathematical analysis*, Undergraduate Texts in Mathematics, Springer-Verlag, Indian Reprint, 2004.