Real Analysis: Measure Theory and Lebesgue Integration.

E. Chernysh

Last updated: May 18, 2017.

Contents

1 Lebesgue Measure in $\mathbb{R}^d$ **2**

1.1 The Lebesgue Outer-Measure . . . 2

1.2 The Lebesgue Measure and the Lebesgue Measure Space . . . 6

1.3 Measurable Functions and Their Sequences . . . 11

1.4 Convergence Theorems . . . 14

2 Lebesgue Theory of Integration **16**

2.1 Simple Functions . . . 16

2.2 Bounded Functions Supported On a Set of Finite Measure . . . 17

2.3 Integration of Non-Negative Measurable Functions . . . 21

2.4 General Integrability and $L^1(\mathbb{R}^d)$ Space . . . 23

2.5 Fubini and Tonelli Theorems . . . 27

3 Differentiation Theorems **32**

3.1 The Hardy-Littlewood Maximal Operator and the Vitali Lemma . . . 32

3.2 The Lebesgue Differentiation Theorems . . . 33

4 Abstract Measure and $L^2(\mathbb{R}^d)$ Space **37**

4.1 Abstract Measure Spaces . . . 37

4.2 Signed Measures and Decompositions . . . 40

4.3 $L^2(\mathbb{R}^d)$ Space . . . 42

Chapter 1

Lebesgue Measure in $\mathbb{R}^d$

1.1 The Lebesgue Outer-Measure

Given any subset $E \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, we can assign it an *outer-measure*. The outer-measure is an extended real value, non-negative, which is obtained by estimating the set E with cubes. This is also sometimes referred to as a *pre-measure*. This is the first step in the development of measure theory in $\mathbb{R}^d$ along with abstract integration theory. The reason we start with an outer-measure is simply because this outer measure will exist for all sets in $\mathbb{R}^d$. However, in some cases it may not behave as we would like it to, and so we shall later define *measurable sets*. The general idea in this section is to define the outer-measure, denoted $\mu^\star$, and study some of the consequential primitive properties.

Definition 1 (Lebesgue Outer-Measure)**.** *Let $E \subseteq \mathbb{R}^d$. The Lebesgue outer-measure, denoted $\mu^\star(E)$, is given by*

$$\mu^\star(E) := \inf_{\{Q_\ell\}_{\ell\in\mathbb{N}}} \left\{ \sum_{\ell\in\mathbb{N}} |Q_\ell| : E \subseteq \bigcup_{\ell\in\mathbb{N}} Q_\ell \right\} \tag{1.1}$$

where $\{Q_\ell\}_{\ell\in\mathbb{N}}$ denotes a countable collection of cubes.

Some remarks are in order:

1. $\mu^\star$ takes value in the non-negative extended real number line: $\overline{\mathbb{R}}_+$. That is, for all sets $E \subseteq \mathbb{R}^d$ one has $0 \leq \mu^\star(E) \leq \infty$.

2. Given a cube Q one has $|Q| = \mu^\star(Q)$ where $|\cdot|$ is our usual notion of volume.

3. We have monotonicity, i.e if $A \subseteq B \subseteq \mathbb{R}^d$ then $\mu^\star(A) \leq \mu^\star(B)$. This is immediate from the observation that any cubic covering of B also covers A. Thus, given any collection of cubes we see that $\mu^\star(A)$ is no greater than the volume of this covering. Since this holds for all such coverings, we may deduce that $\mu^\star(A) \leq \mu^\star(B)$.

Before we proceed let us give some basic, yet essential, results in topology.

Proposition 1.1.1 (Characterization of Open Sets)**.** *Let $G \subseteq \mathbb{R}$ be an open set. Then there exists a disjoint collection of open intervals $(a_n, b_n)_{n\in\mathbb{N}}$ so that*

$$G = \coprod_{n\in\mathbb{N}} (a_, b_n)$$

Proof. Pick a point $x \in G$, there exists δ_x so that the open ball $B(x, \delta_x) \subseteq G$. For any such x pick the largest such ball, this is an interval $x \in (a_x, b_x) \subseteq G$. In this way we define an uncountable collection of open intervals so that

$$G = \bigcup_{x\in G} (a_x, b_x)$$

We claim these intervals are either equal or disjoint. To see this, suppose $(a_x, b_x) \cap (a_y, b_y) \neq \emptyset$ and so we may select a point, say, z from their intersection. If we did not have equality, then we would have, say, $(a_x, b_x) \subset (a_y, b_y) \subseteq G$. But then we could extend this open interval further, contradicting our choice of maximal interval. Therefore, we may assume this collection of intervals is disjoint. To see that they are countable as well, recall that $\mathbb{Q}$ is a dense subset of $\mathbb{R}$ and so from each interval we select some q_x. Since the intervals are disjoint, each rational is unique to each interval. Thus, we may enumerate the intervals by enumeration of the rationals, which we know to be countable. $\square$

Definition 2. *Two cubes are said to be almost disjoint if they are disjoint save, possibly, their boundaries.*

Theorem 1.1.2. *Let $O \subseteq \mathbb{R}^d$ be an open set. We may write O as the union of almost disjoint cubes.*

Proof. Begin by decomposing the open set with grids. As you bisect the interior cubes, each cube becomes smaller and smaller. After each such process, accept the cubes that are completely contained in the cube and proceed recursively with the others. Or, more rigorously: divide the grid into 2^N cubes for any N large. Accept cubes that are already contained in the open set O. For sufficiently large N there will always be such cubes upon the first step, as O is open.

Divide the remaining cubes again into 2^N smaller cubes. Accept those that live inside O and proceed in this way, in countably many steps selecting finitely many cubes in each step. $\square$

Let us now return to our examination of the outer measure.

Proposition 1.1.3 (Sub-Additivity of $\mu^\star$)**.** *Let $E \subseteq \mathbb{R}^d$ and let $\{E_n\}_{n=1}^\infty$ a collection of subsets of $\mathbb{R}^d$ so that $E = \bigcup_{n\in\mathbb{N}} E_n$. Then,*

$$\mu^\star(E) \leq \sum_{n\in\mathbb{N}} \mu^\star(E_n)$$

Proof. For each E_n consider $\mu^\star(E_n)$. First, if any such E_n has $\mu^\star(E) = \infty$ then there is nothing to show. Without loss of generality suppose that $\mu^\star(E_n) < \infty$ for all n. Now let $\varepsilon > 0$ and associate to each E_n a collection $\{Q_{n,j}\}_{j=1}^\infty$ of cubes so that $E_n \subseteq \bigcup_{j=1}^\infty Q_{n,j}$ and $\sum_{j=1}^\infty |Q_{n,j}| \leq \mu^\star(E_n) + \frac{\varepsilon}{2^n}$, the fact that this can always be done is a consequence of the definition of $\mu^\star$. Then, observe that

$$E \subseteq \bigcup_{n,j=1}^\infty Q_{n,j}$$

whence, by definition of the outer-measure:

$$\begin{aligned}\mu^\star(E) \leq \sum_{n,j=1}^\infty |Q_{n,j}| = \sum_{n=1}^\infty \left(\sum_{j=1}^\infty |Q_{n,j}| \right) &\leq \sum_{n=1}^\infty \left(\mu^\star(E_n) + \frac{\varepsilon}{2^n} \right) \\ &= \sum_{n=1}^\infty \mu^\star(E_n) + \sum_{n=1}^\infty \frac{\varepsilon}{2^n} \\ &= \sum_{n=1}^\infty \mu^\star(E_n) + \varepsilon\end{aligned}$$

Let now $\varepsilon \to 0$, to recover in the limit that $\mu^\star(E) \leq \displaystyle\sum_{n=1}^\infty \mu^\star(E_n)$ as was asserted. □

Proposition 1.1.4. *Let $E \subseteq \mathbb{R}^d$. Then $\mu^\star(E) = \inf_{O \supseteq E} \mu^\star(O)$ where O is any such open set.*

Proof. Note that if $E \subseteq O$ then by the monotonicity remark we have $\mu^\star(E) \leq \mu^\star(O)$ for all such O and hence $\mu^\star(E) \leq \inf_{O \supseteq E} \mu^\star(O)$. We shall now only prove the reverse inequality.

Let $\varepsilon > 0$ be given, and let $\{Q_j\}_{j=1}^\infty$ be a collection of cubes covering E in their union with

$$\sum_{j=1}^\infty |Q_j| \leq \mu^\star(E) + \varepsilon$$

Now, for each Q_j we expand the cube a little, forming a slightly larger open cube $Q_j^0 \supset Q_j$ with $\left|Q_j^0\right| \leq |Q_j| + \frac{\varepsilon}{2^j}$. Then, of course we also have

$$E \subseteq \bigcup_{j \in \mathbb{N}} Q_j^0$$

Moreover this union on the right is an open set, as the union of open cubes. Moreover, note that if we denote $O := \bigcup_{j\in\mathbb{N}} Q_j^0$ then by the sub-additive property we deduced prior:

$$\mu^\star(O) \le \sum_{j=1}^{\infty} |Q_j^0| \le \sum_{j=1}^{\infty} \left(|Q_j| + \frac{\varepsilon}{2^j}\right) = \sum_{j=1}^{\infty} |Q_j| + \varepsilon$$
$$= \mu^\star(E) + 2\varepsilon$$

Whence,

$$\inf_{O\supseteq E} \mu^\star(O) \le \mu^\star(E) + 2\varepsilon, \quad \forall \varepsilon > 0$$

Let now $\varepsilon \to 0$ to deduce that $\inf_{O\supseteq E} \mu^\star(O) \le \mu^\star(E)$. These two inequalities proves the claim. $\square$

Lemma 1.1.5. *Let $E_1, E_2 \subseteq \mathbb{R}^d$ and suppose that $d(E_1, E_2) > 0$. Then, $\mu^\star(E_1 \cup E_2) = \mu^\star(E_1) + \mu^\star(E_2)$ where the distance is induced by:*

$$\inf\{|x-y| : x \in E_1, y \in E_2\} \tag{1.2}$$

Proof. Again, by sub-additivity we have already that $\mu^\star(E_1 \cup E_2) \le \mu^\star(E_1) + \mu^\star(E_2)$. We shall again show the reverse inequality. Let now $\{Q_n\}_{n=1}^{\infty}$ be a collection of cubes so that $E = E_1 \cup E_2 \subseteq \bigcup_{n=1}^{\infty} Q_n$ and also $\sum_{n=1}^{\infty} |Q_n| \le \mu^\star(E) + \varepsilon$. Now, we shall use that $d(E_1, E_2) = \delta > 0$. By breaking apart the cubes, i.e partitioning each cube into smaller sub-cubes whilst preserving the overall cubic structure and volume, we may suppose without harm to the proof that each cube has sides strictly less than δ. Now, we observe that no cube Q_j can intersect both E_1 and E_2 simultaneously, for then this would contradict $d(E_1, E_2) = \delta$. Hence, we may extract two completely disjoint sub-collections of cubes $\{Q_j\}_{j\in\mathbb{N}}$ and $\{Q_m\}_{m\in\mathbb{N}}$ so that

$$E_1 \subseteq \bigcup_{j\in\mathbb{N}} Q_j, \quad E_2 \subseteq \bigcup_{m\in\mathbb{N}} Q_m$$

and $\bigcup_n Q_n = \bigcup_j Q_j \cup \bigcup_m Q_m$. Then, we shall then have by the simple fact that these cover E_1 and E_2 respectively:

$$\mu^\star(E_1) + \mu^\star(E_2) \le \sum_j |Q_j| + \sum_m |Q_m| = \sum_n |Q_n| \le \mu^\star(E) + \varepsilon$$

now take $\varepsilon \to 0$. $\square$

Proposition 1.1.6. *If $E \subseteq \mathbb{R}^d$ is the countable union of almost-disjoint cubes: $E = \bigcup_{n=1}^{\infty} Q_n$ then*

$$\mu^\star(E) = \sum_{n=1}^{\infty} |Q_n|$$

Proof. Again, we may assume by monotonicity that if some Q_n has infinite volume, then the equality holds. So assume $|Q_n| < \infty$ for all indices n. Moreover, by sub-additivity we already have $\mu^\star(E) \leq \sum_{n\in\mathbb{N}} |Q_n|$. So, we shall instead show the reverse inequality. Let $N \in \mathbb{N}$ be arbitrary, and $\varepsilon > 0$ freely given. For each Q_j with $j = 1, 2, \cdots, N$ we take a slightly smaller cube $Q_j^0 \subset Q_j$ with $|Q_j| - \frac{\varepsilon}{2^j} \leq \left|Q_j^0\right|$. Then, the collection $\{Q_j^0\}_{j=1}^N \subseteq \{Q_n\}_{n\in\mathbb{N}}$ are disjoint cubes and hence we have that $d(Q_i^0, Q_j^0) > 0$ for all such $i, j = 1, 2, \cdots, N$. Successive applications of the previous lemma yield:

$$\begin{aligned}\mu^\star(E) \geq \mu^\star\left(\bigcup_{j=1}^N Q_j^0\right) = \sum_{j=1}^N |Q_j^0| &\geq \sum_{j=1}^N \left(|Q_j| - \frac{\varepsilon}{2^j}\right) \\ &\geq \sum_{j=1}^N |Q_j| - \varepsilon\end{aligned}$$

Taking $N \to \infty$ we find that in the limit:

$$\mu^\star(E) \geq \sum_{n=1}^\infty |Q_j| - \varepsilon$$

for $\varepsilon > 0$ arbitrary. Letting $\varepsilon \to 0$ finally yields the desired inequality: $\mu^\star(E) \geq \sum_{n=1}^\infty |Q_j|$. □

We now have enough tools to begin our construction of the Lebesgue measure.

1.2 The Lebesgue Measure and the Lebesgue Measure Space

In this section we construct the Lebesgue measure and define the notion of a measurable set. We will study properties of measurable sets and *verify* that these sets obey the expected properties of measure. In fact, we will define several important abstract structures of an analytic nature. We define a measure, σ-algebra, measurable and measure spaces. We will mostly show that the Lebesgue measure is indeed a measure and we shall show that $\mathbb{R}^d$ endowed with the Lebesgue measure is a measure space.

We begin with the outer-measure $\mu^\star$.

Definition 3. *Let $E \subseteq \mathbb{R}^d$. We say that E is **Lebesgue measurable** or simply measurable if and only if for all $\varepsilon > 0$ there exists an open set $\mathcal{O} \supseteq E$ satisfying $\mu^\star(\mathcal{O} \setminus E) \leq \varepsilon$. Whenever this holds, we shall write $\mu(E) := \mu^\star(E)$. This function on sets μ is called the Lebesgue measure.*

Before we proceed, let us make some remarks. First, note that any open set O is necessarily measurable. Indeed, if O is open then it is itself an open-set covering O and of course $\mu^\star(O \setminus O) = \mu^\star(\emptyset) = 0 < \varepsilon$ for all $\varepsilon > 0$. There are also special classes of sets called *null sets*. A set $\mathcal{Z} \subseteq \mathbb{R}^d$ is said to be a *null set* if it has $\mu^\star(\mathcal{Z}) = 0$. These will be very important in the subsequent integration theory, and is even important in the Riemann theory of integration. A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if the set of discontinuity points of f on $[a, b]$ is a null set.

We claim any null set $\mathcal{Z}$ is measurable. Certainly, since it has $\mu^\star(\mathcal{Z}) = 0$ given $\varepsilon > 0$ there exists an open set $\mathcal{O} \supseteq \mathcal{Z}$ so that $\mu^\star(\mathcal{O}) = \mu(\mathcal{O}) < \varepsilon$; by Proposition (1.1.4). But then, $\mathcal{O} \setminus \mathcal{Z} \subseteq \mathcal{O}$ so that $\mu^\star(\mathcal{O} \setminus \mathcal{Z}) \leq \mu^\star(\mathcal{O}) < \varepsilon$. Moreover, a subset of a null set is obviously null by monotonicity and hence is also measurable.

Proposition 1.2.1. *The countable union of measurable sets is also measurable.*

Proof. Let $\{E_n\}_{n\in\mathbb{N}}$ be a collection of measurable sets and set $E := \bigcup_{n=1}^\infty E_n$. Fix $\varepsilon > 0$, and for each E_n select an open set $O_n \supseteq E_n$ with $\mu^\star(O_n \setminus E_n) \leq \varepsilon/2^n$. Consider the open set $O := \bigcup_{n\in\mathbb{N}} O_n$ which clearly contains the set E. Moreover, it is not difficult to see that

$$O \setminus E = \left(\bigcup_{n=1}^\infty O_n\right) \setminus \left(\bigcup_{n=1}^\infty E_n\right) \subseteq \bigcup_{n\in\mathbb{N}} O_n \setminus E_n$$

Whence,

$$\mu^\star(O \setminus E) \leq \sum_{n=1}^\infty \mu^\star(O_n \setminus E_n) \leq \sum_{n=1}^\infty \frac{\varepsilon}{2^n} = \varepsilon \tag{1.3}$$

which proves that $E = \bigcup_n E_n$ is measurable. □

We now seek to extend the class of sets that we know to be measurable. Naturally, open sets are measurable so shouldn't their complements (closed sets) also be measurable? Certainly, we begin with the following useful lemmata:

Lemma 1.2.2. *Let (X, d) be a metric space and $F, K \subseteq X$ with F closed and K compact. If $F \cap K = \emptyset$ then $d(F, K) = \delta > 0$ for some $\delta > 0$.*

Proof. We shall argue by contradiction. Suppose instead that $d(F, K) = 0$. That is,

$$\inf \{d(x, y) : x \in K, y \in F\} = 0$$

So for any $n \in \mathbb{N}$ we may find a pair $(x_n, y_n) \in K \times F$ so that $d(x_n, y_n) \leq \frac{1}{n}$. Now, as (x_n) lives in a compact subspace we may extract a convergent subsequence with a limit $x \in K$. By the above, it is not difficult to see that $y_n \to x$ as $n \to \infty$ too. Since F is closed we must have $x \in F$ and hence $K \cap F \neq \emptyset$: contradiction. □

Lemma 1.2.3. *Compact sets are measurable.*

Proof. Let $\varepsilon > 0$ be given. By compactness, we must have that $\mu^\star(E) < \infty$ for any compact subset of $\mathbb{R}^d$ E. An application of Proposition (1.1.4) guarantees the existence of an open set $O \supseteq E$ with $\mu^\star(O) \leq \mu^\star(E) + \varepsilon$. Moreover, it is clear that $O \setminus E = O \cap E^c$ is open and, therefore, measurable. In other words, by Theorem (1.1.2) we may find a countable collection of almost disjoint cubes $\{Q_j\}_{j=1}^\infty$ so that

$$O \setminus F = \bigcup_{j=1}^\infty Q_j$$

where the interiors of Q_j and Q_i are disjoint for all $i \neq j$. If we fix an index $N \in \mathbb{N}$ then the union $K := \bigcup_{j=1}^N Q_j$ is compact, and moreover is disjoint from F because $E \cap (O \setminus E) = \emptyset$. By our lemma above it follows necessarily that $d(E, K) = \delta > 0$. Moreover, we clearly have since $K \subseteq O$ and $E \subseteq O$ we must have $K \cup E \subseteq O$. Thence,

$$\mu^\star(O) \geq \mu^\star(K \cup E) = \mu^\star(K) + \mu^\star(E) = \sum_{j=1}^N |Q_j| + \mu^\star(E)$$

we may now let $N \to \infty$ to conclude

$$\mu^\star(O) \geq \sum_{j=1}^\infty |Q_j| + \mu^\star(E) = \mu^\star(O \setminus F) + \mu^\star(E)$$

Implying that

$$\mu^\star(O \setminus F) \leq \mu^\star(O) - \mu^\star(E) \leq \varepsilon$$

□

Corollary 1.2.4. *Closed sets are measurable.*

Proof. Let $F \subseteq \mathbb{R}^d$ be closed. This set may equivalently be expressed as the union of compact sets: $F = \bigcup_{n \in \mathbb{N}} F \cap \overline{B(0, n)}$. We know this to be measurable by Proposition (1.2.1). □

Proposition 1.2.5. *If E is measurable then $E^c := \mathbb{R}^d \setminus E$ is also measurable.*

Proof. Suppose E is measurable. Then, given $\varepsilon > 0$ there exists an open set $O_\varepsilon \supseteq E$ so that $\mu^\star(O_\varepsilon \setminus E) \leq \varepsilon$. In other terms, we may construct a sequence of open sets O_n containing E for each index n and $\mu^\star(O_n \setminus E) \leq 1/n$. Define now the set

$$S := \bigcup_{n=1}^\infty O_n^c$$

Each O_n^c is closed, and thus measurable. Again, we conclude from Proposition (1.2.1) that S is measurable as well. Moreover, we note that $S \subseteq E^c$. Similarly,

$$E^c \setminus S \subseteq O_n \setminus E$$

for all n. To see this, observe that if $x \in E^c$ and $x \notin S$ then $x \notin E4$ and $x \notin O_n^c$ for all indices n. My monotonicity we have that $\mu^\star(E^c \setminus S) \leq 1/n$ for all n and thus $\mu^\star(E^c \setminus S) = 0$ whence $E^c \setminus S$ is a null set and consequently is measurable by our previous observations. We now write

$$E^c = S \cup (E^c \setminus S)$$

to see that E^c is measurable. □

Corollary 1.2.6. *Countable intersections of measurable sets are measurable.*

Proof. Let $\{E_n\}_{n\in\mathbb{N}}$ be a countable collection of measurable sets. By the previous proposition we have E_n^c measurable for all n. Then,

$$\left(\bigcap_{n=1}^{\infty} E_n\right)^c = \bigcup_{n=1}^{\infty} E_n^c \tag{1.4}$$

is measurable by Proposition (1.2.1). Applying again Proposition (1.2.5) yields that $\bigcap_{n=1}^{\infty} E_n$ is measurable. □

Theorem 1.2.7. *Let $\{E_n\}_{n=1}^{\infty}$ be a collection of pairwise disjoint measurable sets. Then, additivity holds true:*

$$\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mu(E_n) \tag{1.5}$$

Proof. Suppose without loss of generality that all of the E_n's are bounded. In any case, we note that E_n^c is measurable for all indices n. Thus, for each $\varepsilon > 0$ we may select an open set $O_n \supseteq E_n$ with $\mu^\star(O_n \setminus E_n^c) \leq \varepsilon/2^n$. Consider the set $F_n := O_n^c$, then we claim $E_n \setminus F_n \subseteq O_n \setminus E_n^c$. Now, it follows that $\mu^\star(E_n \setminus F_n) \leq \varepsilon/2^n$ and F_n is closed. Moreover, each $F_n \subseteq E_n$ and hence is compact. Therefore, for $N \in \mathbb{N}$ the union of sets $\bigcup_{n=1}^{N} F_n$ is compact. So, by our previous remarks and propositions:

$$\mu(E) \geq \mu\left(\bigcup_{n=1}^{N} F_n\right) = \sum_{n=1}^{N} \mu(F_n)$$

Now, if A, B are measurable sets then $B = (B \setminus A) \cup A$ implying that $\mu(B) \leq \mu(A) + \mu(B \setminus A)$ or $\mu(B) - \mu(A) \leq \mu(B \setminus A)$. Using this with what we have found above one can see that

$$\mu(E) \geq \sum_{n=1}^{N} \mu(F_n) \geq \sum_{n=1}^{N} (\mu(E_n) - \mu(E_n \setminus F_n)) \geq \sum_{n=1}^{N} \left(\mu(E_n) - \frac{\varepsilon}{2^n} \right) \tag{1.6}$$

$$\geq \sum_{n=1}^{N} \mu(E_n) - \varepsilon \xrightarrow{N \to \infty} \sum_{n=1}^{\infty} \mu(E_n) - \varepsilon \tag{1.7}$$

letting $\varepsilon \to 0^+$ yields the result that we seek. $\square$

Now we turn our attention to a few abstract structures, to study the fruit of our labour and see why all this was important. We begin by defining σ-algebras and measurable spaces. Then we define a measure and auxiliary measure spaces.

Definition 4 (σ-Algebras)**.** *Let X be a set. A collection of subsets of X, Σ, is called a σ-algebra on X provided each of the following hold:*

1. *$X \in \Sigma$.*
2. *Σ is closed under-complement: i.e if $E \in \Sigma$ then $E^c \in \Sigma$.*
3. *Σ is closed under countable union: i.e If $\{E_n\}_{n \in \mathbb{N}} \subseteq \Sigma$ then $\bigcup_{n \in \mathbb{N}} E_n \in \Sigma$.*

Definition 5 (Measurable Spaces)**.** *Let X be a set and Σ a σ-algebra on X. Then the pair (X, Σ) is called a measurable space.*

Definition 6 (Measure)**.** *Let (X, Σ) be a measurable space. A mapping (set function) $\mu : \Sigma \to [0, \infty]$ is called a measure if*

1. *$\mu(\emptyset) = 0$.*
2. *If $\{E_n\}_{n \in \mathbb{N}} \subseteq \Sigma$ is a countable collection of disjoint sets, then*

$$\mu\left(\bigcup_{n \in \mathbb{N}} E_n \right) = \sum_{n \in \mathbb{N}} \mu(E_n) \tag{1.8}$$

Definition 7 (Measure Space)**.** *Let (X, Σ) be a measurable space and μ a measure on Σ. The triple (X, Σ, μ) is called a measure space. If $\mu(X) < \infty$ then we call (X, Σ, μ) a σ-finite measure space. In particular, if $\mu(X) = 1$ then (X, Σ, μ) is called a probability space.*

Therefore, we have the following result based upon our previous results:

Theorem 1.2.8. *Let Σ be the set of Lebesgue measurable sets in $\mathbb{R}^d$ and μ the Lebesgue measure. Then, $(\mathbb{R}^d, \Sigma, \mu)$ is a measure space.*

Proposition 1.2.9. *Let (X, Σ, μ) be a σ-finite measure space and suppose all singletons live in Σ. Then, this space has countably many atoms.*

Proof. Let $n \in \mathbb{N}$ be given, and define $X_n := \left\{x \in X : \mu(\{x\}) \geq \frac{1}{n}\right\}$. Obviously, the collection of atoms in X is given by $\bigcup_{n=1}^{\infty} X_n$. Note that for any index n X_n can have at most finitely many elements, for otherwise $\sigma = \mu(E) \geq \sum_{n^\star} \mu(\{x_{n^\star}\}) = \infty$ which is a contradiction. $\square$

1.3 Measurable Functions and Their Sequences

Here we consider functions $f : \mathbb{R}^d \to \overline{\mathbb{R}} = [-\infty, \infty]$. Naturally, we motivate this section with the following class of functions:

Definition 8. *If $E \subseteq \mathbb{R}^d$ is measurable then a function $f : E \to \overline{\mathbb{R}}$ is said to be measurable if for all $a \in \mathbb{R}$ the pre-image:*

$$f^{-1}([-\infty, a)) \tag{1.9}$$

is measurable. We shall often denote this by $\{f < a\}$.

An important thing to observe, at first, is that there are many equivalent definitions of measurability for functions. Note that for all $a \in \mathbb{R}$ $\{f < a\}$ is measurable if and only if $\{f \geq a\}$ is measurable for all a as well, and this is easily seen by taking complements. Moreover, if $\{f < a\}$ is measurable for all $a \in \mathbb{R}$ then so is $\left\{f < a + \frac{1}{n}\right\}$ for all $n \in \mathbb{N}$. We may now take their unions and deduce from Proposition (1.2.1) that $\{f \leq a\} = \bigcup_{n \in \mathbb{N}} \left\{f < a + \frac{1}{n}\right\}$ is measurable. Conversely, if $\{f \leq a\}$ is measurable for all $a \in \mathbb{R}$ then $\{f < a\} = \bigcup_{n \in \mathbb{N}} \left\{f \leq a - \frac{1}{n}\right\}$ must also be measurable. Considering complements yields equivalent characterizations: $\{f \geq a\}$ and $\{f > a\}$ for all real a. Throughout this text we shall use any characterization that is convenient.

Similarly, we could have defined measurability in terms of $\{a \leq f \leq b\}$ or $\{a < f < b\}$ being measurable for all $a, b \in \mathbb{R}$. This is a reprise of the previous arguments and shall not be explicitly stated here.

Suppose that a function f is finite valued on some measurable set $E \subseteq \mathbb{R}^d$. Let O be an open set in $\mathbb{R}$. Then $f^{-1}(O)$ must be measurable, as O may be expressed as the countable union of disjoint open intervals $\{(a_k, b_k)\}_k$ by Proposition (1.1.1). Then, by earlier remarks each of $\{a_k < f < b_k\}$ must be measurable and of course the set $f^{-1}(O) = \bigcup_k \{a_k < f < b_k\}$ is measurable. Conversely, if $f^{-1}(O)$ is measurable for any open set $O \subseteq \mathbb{R}$ then $\{f < a\}$ is measurable for all $a \in \mathbb{R}$.

It is very useful in practice to note that any continuous function f defined on some measurable set $E \subseteq \mathbb{R}^d$ is necessarily measurable since the pre-image of any open set is open by continuity, and we have already shown that open sets are measurable. In fact, if $f : E \subseteq \mathbb{R}^d \to \overline{\mathbb{R}}$ is measurable with φ is a continuous function defined on a set containing $f(E)$ then the composite $\varphi \circ f$ is measurable. Certainly, let $a \in \mathbb{R}$ be given and consider

$\{\varphi \circ f < a\}$. Since $(\varphi \circ f)^{-1}(X) = f^{-1}(\varphi^{-1}(X))$ we see that $\{\varphi < a\}$ is open by continuity and hence since f is measurable the pre-image of this set is measurable.

Let $(f_n)_{n=1}^{\infty}$ be a sequence of measurable functions. Then, $\sup_n f_n(x)$ is a measurable function of x. Indeed, the main trouble here is making the observation that $\{\sup_n f > a\} = \bigcup_{n=1}^{\infty}\{f_n > a\}$. Certainly, if $\{\sup_n f_n > a\}$ then $f_n > a$ for some n implying $\{\sup_n f > a\} \subseteq \bigcup_{n=1}^{\infty}\{f_n > a\}$. To see the reverse inclusion, note that if $f_n > a$ for some n then clearly $\sup_n f_n > a$ and hence we have the equality of sets thereby establishing measurability. We may proceed similarly with the inf property. Fix $a \in \mathbb{R}$, then a similar argument shows that we may write $\{\inf_n f_n < a\} = \bigcup_n \{f_n < a\}$. Writing out

$$\limsup_{n\to\infty} f_n(x) = \inf_k \left\{\sup_{n\geq k} f_n\right\}, \quad \liminf_{n\to\infty} f_n(x) = \sup_k \left\{\inf_{n\geq k} f_n\right\} \tag{1.10}$$

we have similar results for lim sup and lim inf. Hence, when it exists we have

$$\lim_{n\to\infty} f_n(x) = f(x)$$

is measurable since then $\limsup_n f_n(x) = \lim_{n\to\infty} f_n(x) = f(x)$.

Lemma 1.3.1. *Let $f, g : E \subseteq \mathbb{R}^d \to \mathbb{R}$ be two measurable functions and suppose E is measurable. Then, f^k is measurable for all $k \in \mathbb{N}$ and $f+g$, fg are measurable functions.*

Proof. For the first, note that if $2 \mid k$ then of course $\{f^k > a\} = \{f < -\sqrt[n]{a}\} \cup \{f > \sqrt[n]{a}\}$ and if instead $2 \nmid n$ we have the $\{f^k > a\} = \{f > \sqrt[n]{a}\}$.

To see that $f+g$ is measurable, we recall that $\mathbb{Q}$ is dense in $\mathbb{R}$. Note,

$$\begin{aligned} f+g < a \iff f < a-g &\iff \exists q \in \mathbb{Q}^{\times} \text{ such that } f < q < a-g &(1.11)\\ &\iff \exists q \in \mathbb{Q}^{\times} \text{ such that } f < q \wedge g < a-q &(1.12)\end{aligned}$$

Thence,

$$\{f+g<a\} = \bigcup_{q\in\mathbb{Q}}\{f<q\}\cap\{g<a-q\}$$

which is certainly measurable since $\mathbb{Q}$ is countable and our results from the previous section apply. For the product, note that we may write:

$$fg = \frac{1}{2}\left((f+g)^2 - (f-g)^2\right) \tag{1.13}$$

□

We introduce some terminology now. We say that an event occurs **almost everywhere** if the only exception is a null set. For instance, we could say that a sequence of functions f_n converge point-wise to f **almost everywhere**, or a.e, if

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \notin Z$$

for a null set Z (this may be empty). Note that if f is measurable and $f = g$ almost everywhere then g is measurable. Certainly, there exists a set of measure 0, say X, so that $f = g$ for all $x \notin X$. Then, given $a \in \mathbb{R}$ one may express $\{g < a\} = \{x \notin X : g(x) < a\} \sqcup \{x \in X : g(x) < a\}$. The latter set has measure 0, and is hence measurable.

Now, given a set E, not necessarily measurable we define it's *characteristic function*, denoted $\mathbb{1}_E$, by

$$\mathbb{1}_E(x) := \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases} \tag{1.14}$$

We may now give an immediate generalization of step functions. Given a finite sequence of measurable sets $E_1, E_2, \cdots, E_N$ and an associated sequence of constant complex numbers $a_1, a_2, \cdots, a_N$ a simple function is a function of the form:

$$f(x) = \sum_{n=1}^{N} a_n \mathbb{1}_{E_n}(x) \tag{1.15}$$

Observe that step functions are simple functions, as intervals are measurable.

Theorem 1.3.2. *Suppose f is a non-negative measurable function defined on $\mathbb{R}^d$. There exists a sequence of non-negative simple functions, increasing, converging point-wise to $f(x)$. That is,*

$$\varphi_n(x) \leq \varphi_{n+1}(x), \quad \varphi_n(x) \xrightarrow{n\to\infty} f(x) \tag{1.16}$$

for some sequence (φ_n).

We shall omit this very tedious proof.

Corollary 1.3.3. *Suppose that f is measurable on $\mathbb{R}^d$. There exists a sequence of simple functions (φ_n) satisfying*

$$|\varphi_n(x)| \leq |\varphi_{n+1}(x)|, \quad \varphi_n(x) \xrightarrow{n\to\infty} f(x) \tag{1.17}$$

Proof. Write $f^+ := \max(f(x), 0)$ and $f^- = \max(-f(x), 0)$. Clearly,

$$f(x) = f^+(x) - f^-(x), \quad \forall x \in \mathbb{R}^d$$

By the previous theorem, we may take two sequences of simple functions $(\phi_n), (\psi_n)$ monotonically increasing to f^+ and f^- respectively. Define now

$$\varphi_n(x) := \phi_n(x) - \psi_n(x)$$

Clearly, $\varphi_n(x) \to f(x)$ as $n \to \infty$. Moreover, $|\varphi_n(x)| = \phi_n(x) + \psi_n(x)$. $\square$

Theorem 1.3.4. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be measurable. There exists a sequence of step functions (φ_n) converging point-wise to f almost everywhere.*

1.4 Convergence Theorems

Theorem 1.4.1 (Egorov's Theorem)**.** *Let (f_k) be a sequence of measurable functions converging point-wise to a function f almost everywhere, defined on a set E with $\mu(E) < \infty$. For each $\varepsilon > 0$ there exists a closed set $A_\varepsilon \subset E$ so that $\mu(E \backslash A_\varepsilon) \leq \varepsilon$ and $\varphi_n \to f$ uniformly on A_ε.*

Proof. We shall suppose, without loss of generality, that $f_n \to f$ everywhere on E, point-wise. Given n, k define the auxiliary set

$$E_k^n := \left\{ x \in E : |f_j(x) - f(x)| < \frac{1}{n}, \quad \forall j > k \right\}$$

Suppose we fix n; then $E_k^n \subseteq E_{k+1}^n$ for all $k \in \mathbb{N}$. Thus, this is an increasing sequence in k. Moreover, $E_k^n \xrightarrow{k\to\infty} E$. So there exists a $k_n \in \mathbb{N}$ so that $\mu(E \setminus E_{k_n}^n) \leq 2^{-n}$. Now, let N be so large that $\sum_{n \geq N} 2^{-n} \leq \varepsilon$. Consider now the set

$$F := \bigcap_{n \geq N} E_{k_n}^n$$

We claim that $f_k \xrightarrow{k\to\infty} f$ uniformly on F. Indeed, for $\gamma > 0$, take $n \geq N$ so large that $1/n \leq \gamma$. If $x \in F$, then we have that $|f_j(x) - f(x)| < \gamma$ for all $j > n$: which shows that the convergence is uniform. As for the measure of this set, note that

$$\begin{aligned}
\mu\left(E \setminus F\right) &= \mu\left(E \setminus \bigcap_{n\geq N} E_{k_n}^n\right) = \mu\left(E \cap \bigcup_{n \geq N} \left(E_{k_n}^n\right)^c\right) \\
&= \mu\left(\bigcup_{n\geq N} E \cap \left(E_{k_n}^n\right)^c\right) = \mu\left(\bigcup_{n\geq N} E \setminus E_{k_n}^n\right) \\
&\leq \sum_{n\geq N} \mu(E \setminus E_{k_n}^n) \leq \sum_{n \geq N} 2^{-n} \leq \varepsilon
\end{aligned}$$

It now suffices to take a closed subset with $\mu(F \setminus A_\varepsilon)$ small. □

Theorem 1.4.2 (Lusin's Theorem)**.** *Let f be measurable and and defined on some set $E \subseteq \mathbb{R}^d$ with $\mu(E) < \infty$ and $f(x) < \infty$ everywhere. For every $\varepsilon > 0$, there exists a closed set $A_\varepsilon \subseteq E$ with $\mu(E \setminus A_\varepsilon) \leq \varepsilon$ and $f_{|A_\varepsilon}$ continuous.*

Proof. Let $\varepsilon > 0$ be given and consider a sequence (φ_n) of step functions converging point-wise to $f(x)$, save possibly a set of measure 0. In any case, these step functions are continuous save a set of measure 0, X. So, for each n, there exists a set E_n with

$\mu(E_n) \leq 2^{-n}$ and φ_n continuous outside of E_n. Moreover, by Egorov's theorem we may select a set A_ε so that $\mu(E \setminus F_\varepsilon) \leq \varepsilon$ and $\varphi_n \to f$ uniformly on F_ε. Now, define the set

$$\widetilde{A_\varepsilon} := F_\varepsilon \setminus \bigcup_{n \geq N} E_n$$

for N so large that $\sum_{n \geq N} 2^{-n} \leq \varepsilon$. Then, all φ_n are continuous on $\widetilde{A_\varepsilon}$ and as the uniform limit of continuous functions so is f. Now,

$$\begin{aligned} \mu\left(E \setminus \widetilde{A_\varepsilon}\right) &= \mu\left[E \setminus \left(F_\varepsilon \setminus \bigcup_{n \in \mathbb{N}} E_n\right)\right] \\ &\leq \sum_{n \geq N} \mu(E_n) \leq \varepsilon \end{aligned}$$

Now, approximate by a smaller closed subset as before. $\square$

Chapter 2

Lebesgue Theory of Integration

In this chapter we study the Lebesgue integral in $\mathbb{R}^d$, for a natural number d. It should become obvious to the reader that this is a tremendous improvement over the Riemann integral. A somewhat old, but difficult problem, was to give a sufficient and necessary condition for the Riemann integrability of a function. Lebesgue answered this question, but to do so he had to create an entire new theory of integration. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if

$$\mathscr{Z} := \{x \in [a, b] : f(x) \text{ is discontinuous at } x\}$$

is a null set. Here we shall construct this integral, in $\mathbb{R}^d$, although many of the methods will generalize immediately to an arbitrary measure space (X, Σ, μ). We give this in $\mathbb{R}^d$ rather because it is most applicable and intuitive. Although, most of the proofs will be just as valid in a measure space.

2.1 Simple Functions

In this short section we define the Lebesgue integral of a simple function, which was defined in the previous section. Recall that a simple function is a function of the form

$$f(x) = \sum_{k=1}^{N} a_k \mathbb{1}_{E_k}(x), \quad a_k \in \mathbb{R}$$

and the E_k's measurable. We shall instead only consider the **canonical** forms of these simple functions. Given a simple function $\psi(x)$, the we say $\sum_{k=1}^{N} a_k \mathbb{1}_{E_k}(x)$ is the canonical form of ψ provided all the a_k's are distinct and the E_k's are disjoint. In this case, one defines:

$$\int_{\mathbb{R}^d} \psi = \int_{\mathbb{R}^d} \psi(x)\, \mathrm{d}x := \sum_{k=1}^{N} a_k \mu(E_k) \tag{2.1}$$

for the canonical form. If we wish to integrate over a proper measurable subset E of $\mathbb{R}^d$ we define

$$\int_E \psi = \int_E \psi(x)\,\mathrm{d}x := \int_{\mathbb{R}^d} \psi(x)\mathbb{1}_E \tag{2.2}$$

The following are consequences of this definition:

Proposition 2.1.1. *The Lebesgue integral of simple functions satisfy:*

(a) Independence of representation. Any canonical form yields the same integral value.

(b) Linearity: for any pair of simple functions φ, ϕ then, $\int_{\mathbb{R}^d} (\varphi + \phi) = \int_{\mathbb{R}^d} \varphi + \int_{\mathbb{R}^d} \phi$.

(c) If E, F are disjoint measurable sets, then $\int_{E\cup F} \varphi = \int_E \varphi + \int_F \varphi$.

(d) If $\varphi \leq \phi$ then $\int_{\mathbb{R}^d} \varphi \leq \int_{\mathbb{R}^d} \phi$. *In particular, this holds for $\varphi \leq |\varphi|$.*

These are all obvious. The first is due to the fact that if the E_k's are disjoint then we may adjoin sets where the function takes the same complex value, obtaining the canonical form. If instead E_k's are not disjoint then we can break them apart and repeat the process above. For the linearity, we need only notice that the summation is linear. When considering (c), simply observe that $\mathbb{1}_{E\cup F} = \mathbb{1}_E + \mathbb{1}_F$ since $E \cap F = \emptyset$.

2.2 Bounded Functions Supported On a Set of Finite Measure

We now write $\mathrm{supp}(f)$ to denote the support of f, given any function f. More precisely, if X is a set and V a vector space, the **support** of a function $f : X \to V$, denoted $\mathrm{supp}(f)$, is the set

$$\mathrm{supp}(f) := \{x \in X : f(x) \neq \mathbf{0}_V\} \tag{supp}$$

In other terms, $\mathrm{supp}(f) = X \setminus \ker f$. Note that if f is continuous with compact support over a metric space, then it is necessarily uniformly continuous over the entire space. Throughout this section we shall study the integration theory of measurable functions f with $\mu(\mathrm{supp}(f)) < \infty$. Note that writing:

$$\mathrm{supp}(f) = \{f < 0\} \cup \{f > 0\}$$

we have that $\mathrm{supp}(f)$ is measurable whenever f is a measurable function. Now, we define the Lebesgue integral of such a function by:

$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi_n(x)\,\mathrm{d}x \tag{2.3}$$

where (φ_n) is a sequence of simple functions converging point-wise to f almost everywhere, bounded by the same constant as f; which we know to exist by the previous chapter. We extend this to an integral over arbitrary measurable sets by attaching a factor of the indicator function, just as we did in the last section for simple functions.

Several things are in order. We need to show that this is well-defined, i.e that it is independent of the choice of (φ_n). In fact, we need to verify that this limit will always exist, as this is certainly not obvious. Then we shall show the independence of sequence. Let us begin with the following lemma:

Lemma 2.2.1. *Let f be measurable on $\mathbb{R}^d$ and suppose that it is bounded, and supported on a set of finite measure E. Then, if $\varphi_n \to f$ is a sequence of simple functions converging point-wise, bounded by the same constant as f, the limit*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi_n(x)\,\mathrm{d}x$$

exists.

Proof. Recall that $\mathbb{R}$ is complete, so is suffices to show that the sequence of integrals is Cauchy. Certainly, let $\varepsilon > 0$ be given. By Egorov's theorem, there exists a closed set $F \subset E$ with $\mu(E \setminus F) \leq \varepsilon$ so that $\varphi_n \to f$ uniformly on E. In this case, since the convergence is uniform and there is a bound M on the convergent sequence we find for all large n, m:

$$\begin{aligned}
\left| \int_{\mathbb{R}^d} \varphi_n - \int_{\mathbb{R}^d} \varphi_m \right| &\leq \int_{\mathbb{R}^d} |\varphi_n - \varphi_m| = \int_E |\varphi_n - \varphi_m| \\
&= \int_F |\varphi_n - \varphi_m| + \int_{E\setminus F} |\varphi_n - \varphi_m| \\
&\leq \int_F \varepsilon + 2M\mu(E \setminus F) \\
&\leq \varepsilon\mu(E) + 2M\varepsilon
\end{aligned}$$

which proves the claim. □

In fact, this same argument proves that if $f = 0$ almost everywhere, then

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \varphi_n(x)\,\mathrm{d}x = 0$$

Certainly, for $\varepsilon > 0$ we may write for all $n \gg 0$:

$$\begin{aligned}
\left| \int_E \varphi_n(x)\,\mathrm{d}x \right| \leq \int_E |\varphi_n(x)|\,\mathrm{d}x &= \int_F |\varphi_n(x)|\,\mathrm{d}x + \int_{E\setminus F} |\varphi_n(x)|\,\mathrm{d}x \\
&\leq \varepsilon\mu(E) + 2M\varepsilon
\end{aligned}$$

It remains to show that this definition is independent of representation. Suppose we have two sequences of equi-bounded simple functions, $(\varphi_n), (\phi_n)$ converging point-wise to f almost-everywhere. Consider the auxiliary sequence $\psi_n := \varphi_n - \phi_n$. It is easy to see that $\psi_n \to 0$ almost everywhere. Moreover, it is certainly a bounded simple function. Therefore, the previous remark implies that

$$\lim_{n\to\infty} \int_E (\varphi_n(x) - \phi_n(x))\,\mathrm{d}x = 0$$

which shows the limits are equal, and hence that the integral is well-defined. Note that the usual integral properties presented in the previous section carry over to this class of functions, as we have defined the integral in terms of a limit of integrals of simple functions, which behave as expected. We have the following result:

Theorem 2.2.2 (Bounded Convergence Theorem)**.** *Let (f_n) be a sequence measurable functions supported on a set E with $\mu(E) < \infty$ and suppose the f_n's are bounded by a constant $M > 0$ in $\mathbb{R}^d$. If $f_n(x)$ converges point-wise to a function f, then*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x \tag{2.4}$$

Proof. The limit $f(x)$ is measurable by previous results. Let $\varepsilon > 0$ be given. There exists a closed set $F \subset E$ with $\mu(E \setminus F) \leq \varepsilon$ and $f_n \to f$ uniformly on F by Egorov's theorem. Then, for all sufficiently large n:

$$\begin{aligned}\left| \int_{\mathbb{R}^d} f_n - \int_{\mathbb{R}^d} f \right| \leq \int_{\mathbb{R}^d} |f_n - f| &\leq \int_F |f_n - f| + \int_{E\setminus F} |f_n - f| \\ &\leq \varepsilon\mu(E) + 2M\varepsilon\end{aligned}$$

□

Again, if f is measurable with $\mu(\mathrm{supp}(f)) < \infty$ and $f \geq 0$ everywhere, then $\int_{\mathbb{R}^d} f = 0$ implies $f = 0$ almost everywhere. Certainly, define the set $X_n := \{f > 1/n\}$. Now note that

$$\frac{1}{n}\mu(X_n) \leq \int_{X_n} f \leq \int_{\mathbb{R}^d} f = 0 \implies \mu(X_n) = 0$$

Hence, $\mu(X_n) = 0$ for all n and thus $\mu(\mathrm{supp}(f)) = \mu\left(\bigcup_n X_n\right) \leq \sum_n \mu(X_n) = 0$ as was asserted.

For this definition to be of any practical use, it must agree with the Riemann integral, whenever it exists. We prove this below:

Theorem 2.2.3 (Consistency Theorem)**.** *Let f be Riemann integrable. Then f is Lebesgue integrable, and this same integral agrees with the Riemann integral.*

Proof. Clearly f is bounded and is supported on a set of finite measure, namely an interval of the form $[a, b] \subseteq \mathbb{R}$. To prove that f is also Lebesgue integrable, we shall show that f is measurable. Now, since f is Riemann integrable we may select two monotonic increasing sequences of step functions bounded uniformly by a constant $M > 0$ such that

$$\varphi_1(x) \leq \varphi_2(x) \leq \cdots \leq \varphi_n(x) \cdots \leq f(x) \leq \cdots \leq \phi_n(x) \leq \cdots \leq \phi_1(x) \tag{2.5}$$

so that

$$\lim_{n\to\infty} \int_{[a,b]}^{\mathcal{R}} \varphi_n(x)\,\mathrm{d}x = \int_{[a,b]}^{\mathcal{R}} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{[a,b]}^{\mathcal{R}} \phi_n(x)\,\mathrm{d}x$$

By very definition of the Lebesgue integral for simple functions (and hence step functions), we have that

$$\int_{[a,b]}^{\mathcal{R}} \varphi_k(x)\,\mathrm{d}x = \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)\,\mathrm{d}x$$

Now, consider the sequence $\psi_n := \phi_n - \varphi_n$; clearly $\psi_n \geq 0$. Define now

$$\psi(x) := \lim_{n\to\infty} \varphi_n(x), \quad \widetilde{\psi}(x) := \lim_{n\to\infty} \phi_n(x)$$

These limits certainly exist, as the limit of monotonic bounded sequences. Note that $\psi(x) \leq f(x) \leq \widetilde{\psi}(x)$ and so $\widetilde{\psi} - \psi \geq 0$. Moreover, since $\phi_n - \varphi_n$ is a sequence of bounded simple functions converging point-wise to $\widetilde{\psi} - \psi$ we have by very definition:

$$\begin{aligned}\int_{[a,b]}^{\mathcal{L}} \left(\widetilde{\psi} - \psi\right) &= \lim_{n\to\infty} \left(\int_{[a,b]}^{\mathcal{L}} \phi_n(x)\,\mathrm{d}x - \int_{[a,b]}^{\mathcal{L}} \varphi_n(x)\,\mathrm{d}x \right) \\ &= \lim_{n\to\infty} \left(\int_{[a,b]}^{\mathcal{R}} \phi_n(x)\,\mathrm{d}x - \int_{[a,b]}^{\mathcal{R}} \varphi_n(x)\,\mathrm{d}x \right) = 0\end{aligned}$$

By the preceding remark, we must have $\widetilde{\psi} = \psi$ almost everywhere, whence $\psi(x) = f(x) = \widetilde{\psi}(x)$ almost everywhere. Thence, f must be measurable. Now, to see that the integrals agree simply write:

$$\begin{aligned}\int_{[a,b]}^{\mathcal{L}} f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_n(x)\,\mathrm{d}x &= \lim_{n\to\infty} \int_{[a,b]}^{\mathcal{R}} \varphi_n(x)\,\mathrm{d}x \\ &= \int_{[a,b]}^{\mathcal{R}} f(x)\,\mathrm{d}x\end{aligned}$$

which is what we had to show. $\square$

2.3 Integration of Non-Negative Measurable Functions

In this section we seek to generalize the results of the previous section unto non-negative measurable functions. That is, we shall learn to integrate non-negative functions f defined on $\mathbb{R}^d$, with not-necessarily bounded support and possibly infinitely valued.

Suppose f is defined on $\mathbb{R}^d$, with $\mu(\mathrm{supp}(f)) \leq \infty$, and let g denote any measurable function satisfying $0 \leq g \leq f$ everywhere that is bounded and $\mathrm{supp}(g)$ has finite measure. We shall then define the (Lebesgue) integral of f:

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x := \sup_{0\leq g\leq f} \int_{\mathbb{R}^d} g(x)\,\mathrm{d}x \tag{2.6}$$

We shall call f **integrable** if

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x < \infty$$

It is clear that this is well defined, as there is only one sup. Yet again, given a measurable set E we have the natural extension:

$$\int_E f(x)\,\mathrm{d}x := \int_{\mathbb{R}^d} f(x)\mathbb{1}_E(x)\,\mathrm{d}x \tag{2.7}$$

Many of the usual properties carry over, as we need only pass to the sup to preserve them. One of much less obvious, and this is the linearity of the integral. We wish to show that for suitable functions f, g as given: $\int_{\mathbb{R}^d}(af + bg) = a\int_{\mathbb{R}^d} f + b\int_{\mathbb{R}^d} g$. We shall show this before proceeding to other (very important) convergence theorems. If we choose such finitely supported functions $0 \leq \eta_1 \leq f$ and $0 \leq \eta_2 \leq g$ we immediately see that $(a\eta_1 + b\eta_2)$ is a candidate for the integral of $(af + bg)$. We have $\eta_1 \leq f$ and $\eta_2 \leq g$ so that $\eta_1 + \eta_2 \leq (f + g)$. Then,

$$\sup_{\eta_1,\eta_2} \int_{\mathbb{R}^d} (\eta_1 + \eta_2) \leq \int_{\mathbb{R}^d} (f + g)$$

But this left hand side corresponds to $\int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$ and so we have

$$\int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g \leq \int_{\mathbb{R}^d} (f + g)$$

For the reverse inequality, suppose that η is an admissible candidate for $\int_{\mathbb{R}^d}(f+g)$. Consider the function $\eta_1(x) := \min(f(x), \eta(x))$ and $\eta_2(x) := (\eta - \eta_1)(x)$. Note that $\eta_1(x) \leq f(x)$ and so is admissible for the integral of f. Moreover, $\eta_2(x) \leq g(x)$. To see this, distinguish both cases. If $\eta_1(x) = f(x)$ then $\eta_2 = \eta(x) - f(x) \leq g(x)$ and it vanishes if $\eta_1(x) = \eta(x)$. Of course;

$$\int_{\mathbb{R}^d} \eta = \int_{\mathbb{R}^d} \eta_1 + \int_{\mathbb{R}^d} \eta_2 \leq \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g \tag{2.8}$$

so

$$\int_{\mathbb{R}^d} (f+g) = \sup_{\eta} \int_{\mathbb{R}^d} \eta \le \int_{\mathbb{R}^d} f + \int_{\mathbb{R}^d} g$$

which completes the proof.

Lemma 2.3.1 (Fatou's Lemma)**.** *Let (f_n) be a sequence of non-negative measurable functions converging point-wise to a function $f(x)$ for almost all x. Then,*

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x \le \liminf_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x \tag{2.9}$$

Proof. Let g be an admissible function for the integral of f: i.e let $0 \le g \le f$ everywhere with g measurable, bounded, and supported on a set of finite measure. Finally define an auxiliary sequence (g_n) by setting $g_n(x) := \min(g, f_n)$. Since $f_n(x) \to f(x)$ almost everywhere, and $g \le f$ everywhere, it is clear that $g_n(x) \to g(x)$ almost everywhere. Moreover, $g_n \le f_n$ everywhere and for all n, thus

$$\int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x \le \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x$$

passing to the $\liminf$ we have

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x \le \liminf_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x$$

But by *bounded convergence* (Theorem 2.2.2) we have that

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_{\mathbb{R}^d} g_n(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} g(x)\,\mathrm{d}x$$

So, taking the sup over admissible g yields:

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \sup_{g} \int_{\mathbb{R}^d} g(x)\,\mathrm{d}x \le \liminf_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x \tag{2.10}$$

which is the desired result. □

Lemma 2.3.2. *Let f be a non-negative measurable function and suppose there exists a sequence (f_n) of non-negative measurable functions with $f_n(x) \le f(x)$ and $f_n \to f$ for almost all x. Then,*

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x \le \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$$

In fact, invoking Fatou's lemma here will establish the equality of integrals above.

Proof. Since $f_n(x) \leq f(x)$ for each and every index n one has

$$\int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x \leq \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$$

whence, passing to the lim sup:

$$\limsup_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x \leq \limsup_{n\to\infty} \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$$

□

A very powerful corollary follows:

Corollary 2.3.3 (Monotone Convergence Theorem)**.** *Let f be a non-negative measurable function and (f_n) a sequence of non-negative measurable functions with $f_n \leq f_{n+1}$ and suppose that $f_n \to f$ almost everywhere. Then,*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x \tag{2.11}$$

Corollary 2.3.4. *Let $f(x) := \sum_{n=1}^{\infty} a_n(x)$ for measurable, non-negative functions $a_n(x)$. Suppose that $\sum_{n=1}^{\infty} \int_{\mathbb{R}^d} a_n(x)\,\mathrm{d}x < \infty$, then $f(x)$ is integrable and the series converges almost everywhere.*

Proof. Fix $N \gg 0$, consider:

$$\int_{\mathbb{R}^d} \sum_{n=1}^{N} a_n(x)\,\mathrm{d}x = \sum_{n=1}^{N} \int_{\mathbb{R}^d} a_n(x)\,\mathrm{d}x$$

Is is clear that $f_N(x) := \sum_{n=1}^{N} a_n(x) \uparrow f(x)$ as $N \to \infty$ and are non-negative. It follows from the monotone convergence theorem (Theorem (2.3.3)) that

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\mathbb{R}^d} a_n(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} \int_{\mathbb{R}^d} a_n(x)\,\mathrm{d}x < \infty$$

which implies that the series is integrable and convergent almost everywhere. □

2.4 General Integrability and $L^1(\mathbb{R}^d)$ Space

The definition here is easy. If f is measurable, then we shall set

$$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f^+(x)\,\mathrm{d}x - \int_{\mathbb{R}^d} f^-(x)\,\mathrm{d}x \tag{2.12}$$

Again, we shall call f **integrable** whenever $\int_{\mathbb{R}^d} |f| < \infty$. The reader should take note that this is an immense improvement over the theory of Riemann. His integral required continuity almost everywhere along with boundedness, this integral requires only measurability and being finite save possibly a null set. As the sum of two integrals, the previous properties obviously carry over unto this general case.

Naturally, this induces a space. We define

Definition 9. *$L^1(\mathbb{R}^d)$ is the space of all functions defined on a subset of $\mathbb{R}^d$ such that $\displaystyle\int_{\mathbb{R}^d} |f(x)| \, \mathrm{d}x < \infty$.*

We can define a norm on this space. Consider:

$$\|\cdot\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\cdot| \, \mathrm{d}x \tag{2.13}$$

It is left to the reader to verify that the pair $(L^1(\mathbb{R}^d), \|\cdot\|_{L^1(\mathbb{R}^d)})$ induces a metric space. We shall, however, show that this metric space is complete; later on.

Lemma 2.4.1. *Let $f \in L^1(\mathbb{R}^d)$ be given. If $\varepsilon > 0$ is given, then there exists a set of finite measure B, so that*

$$\int_{B^c} |f(x)| \, \mathrm{d}x < \varepsilon$$

Proof. Let $\varepsilon > 0$ be fixed. Without loss of generality we may suppose that f is non-negative, replacing f with $|f|$ otherwise. Define a sequence of functions $f_N(x) := f(x)\mathbb{1}_{B(0,N)}(x)$. It should be clear to the reader that these f_N increase to the function $f(x)$. Moreover, it is clear that each such f_N need be measurable. Then, by the monotone convergence theorem (Theorem 2.3.3) we have

$$\lim_{N\to\infty} \int_{\mathbb{R}^d} f_N(x) \, \mathrm{d}x = \int_{\mathbb{R}^d} f(x) \, \mathrm{d}x$$

Thus, for N sufficiently large,

$$\int_{\mathbb{R}^d} f(x) \, \mathrm{d}x - \int_{B(0,N)} f(x) \, \mathrm{d}x = \int_{B_N^c} f(x) \, \mathrm{d}x < \varepsilon \tag{2.14}$$

□

Lemma 2.4.2. *Let f be Lebesgue integrable (i.e $f \in L^1(\mathbb{R}^d)$). Then, for each $\varepsilon > 0$, there exists $\delta > 0$ so that whenever $\mu(E) < \delta$ then*

$$\int_E |f(x)| \, \mathrm{d}x < \varepsilon \tag{2.15}$$

Proof. Without harm assume f is non-negative. Let $\varepsilon > 0$ be given. Define for $N \in \mathbb{N}$ the sequence of functions $f_N := f(x)\mathbb{1}_{E_N}$ where $E_N := \{|f| \leq N\}$. Clearly, $f_N \uparrow f$ as $N \to \infty$. So, by the monotone convergence theorem we have for large enough N: and any measurable set E:

$$\begin{aligned}\int_E f = \int_E (f - f_N) + \int_E f_N &\leq \int_{\mathbb{R}^d} (f - f_N) + N\mu(E) \\ &\leq \varepsilon + N\mu(E) \\ &< \varepsilon\end{aligned}$$

if $\mu(E) < \delta$ for sufficiently small δ. □

We are now sufficiently equipped to give the following powerful result:

Theorem 2.4.3. *Suppose (f_n) are a sequence of measurable functions converging to a function f almost everywhere and suppose $|f_n| \leq g$, where g is an integrable function. Then,*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x \tag{2.16}$$

Proof. Let $\varepsilon > 0$, and given $N \in \mathbb{N}$ define the set

$$E_N := \{x : |x| \leq N \text{ and } |g(x)| \leq N\}$$

By the previous lemmata, and the integrability of g we may find N so large that

$$\int_{E_N^c} |g(x)|\,\mathrm{d}x < \varepsilon$$

Now, we note that

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x - \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x\right| &\leq \int_{\mathbb{R}^d} |f - f_n|(x)\,\mathrm{d}x \\ = \int_{E_N} |f - f_n|(x)\,\mathrm{d}x &+ \int_{E_N^c} |f - f_n|(x)\,\mathrm{d}x\end{aligned}$$

Now, by Theorem (2.2.2) we may take n large enough so that $\int_{E_N} |f - f_n|(x)\,\mathrm{d}x < \varepsilon$. Then, we see that the above is bounded above by $\varepsilon + 2\varepsilon$. That is,

$$\left|\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x - \int_{\mathbb{R}^d} f_n(x)\,\mathrm{d}x\right| \leq \varepsilon + 2\varepsilon = 3\varepsilon$$

which concludes the proof. □

Lemma 2.4.4. *Let (X, d) be a metric space and (x_n) a Cauchy sequence in this space. If (x_n) has a convergent subsequence, the entire sequence converges,* *Proof.* Let $\varepsilon > 0$ be given, there exists a subsequence (x_{n_k}) with a limit point, say, $x \in X$. For such $\varepsilon > 0$, we may find $N \in \mathbb{N}$ so that for all $n_k \geq N$: $d(x_{n_k}, x) < \varepsilon$. Since this sequence is also Cauchy, for this same ε there exists $N' \in \mathbb{N}$ so that for any pair $(n,m) \in \mathbb{N} \times \mathbb{N}$ with $n, m \geq N'$ one has $d(x_n, x_m) < \varepsilon$. Now set $\widetilde{N} := \max(N, N^prime)$ and let $n_k > N$. Then, for all $n \geq N$:

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) \leq 2\varepsilon$$

$\square$

Theorem 2.4.5. *The space* $(L^1(\mathbb{R}^d), \|\cdot\|_{L^1(\mathbb{R}^d)})$ *is complete.*

Proof. Let $(f_n) \subseteq L^1(\mathbb{R}^d)$ be a Cauchy sequence. We shall extract a convergent subsequence. Since this sequence is Cauchy, we may find a subsequence (f_{n_k}) so that $\left\|f_{n_{k+1}} - f_{n_k}\right\|_{L^1(\mathbb{R}^d)} < 2^{-k}$. If this is not obvious, for each k take $\varepsilon := 2^{-k}$ and take the associated N. For the next k, choose an associated N larger than the previous and so on.

Now that we have this subsequence, we shall show that it is convergent in this space. First, consider the function:

$$f(x) := f_{n_1}(x) + \sum_{k=1}^{\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \tag{2.17}$$

Note that the partial sums of the series defining f are precisely the f_{n_k}, which live in our space. We shall now show that f also belongs to the space $L^1(\mathbb{R}^d)$. Certainly, we write:

$$|f| \leq |f_{n_1}|(x) + \sum_{k=1}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right|$$

whence by Corollary (2.3.4)

$$\int_{\mathbb{R}^d} |f(x)| \,\mathrm{d}x \leq \int_{\mathbb{R}^d} |f_{n_1}(x)| \,\mathrm{d}x + \int_{\mathbb{R}^d} \sum_{k=1}^{\infty} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right| \,\mathrm{d}x \tag{2.18}$$

$$= \|f_{n_1}\|_{L^1(\mathbb{R}^d)} + \sum_{k=1}^{\infty} \int_{\mathbb{R}^d} \left|f_{n_{k+1}}(x) - f_{n_k}(x)\right| \,\mathrm{d}x \tag{2.19}$$

$$= \|f_{n_1}\|_{L^1(\mathbb{R}^d)} + \sum_{k=1}^{\infty} \left\|f_{n_{k+1}} - f_{n_k}\right\|_{L^1(\mathbb{R}^d)} \tag{2.20}$$

$$\leq \|f_{n_1}\|_{L^1(\mathbb{R}^d)} + \sum_{k=1}^{\infty} \frac{1}{2^k} = \|f_{n_1}\|_{L^1(\mathbb{R}^d)} + 1 < \infty \tag{2.21}$$

It remains to show that $(f_{n_k}) \to f$ in $L^1(\mathbb{R}^d)$. Indeed, by Fatou's lemma:

$$\begin{aligned}\|f_{n_k} - f\|_{L^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |f_{n_k} - f| &= \int_{\mathbb{R}^d} \lim_{\ell\to\infty} |f_{n_k} - f_{n_\ell}| \\ &\le \liminf_{\ell\to\infty} \int_{\mathbb{R}^d} |f_{n_k} - f_{n_\ell}| \\ &= \liminf_{\ell\to\infty} \|f_{n_k} - f_{n_\ell}\|_{L^1(\mathbb{R}^d)}\end{aligned}$$

But for k very large this last term can be made arbitrarily small, since the sequence is Cauchy. This completes the proof of the theorem. We have shown that $(L^1(\mathbb{R}^d), \|\cdot\|_{L^1(\mathbb{R}^d)})$ is complete. $\square$

There are several nice things to be said about $L^1(\mathbb{R}^d)$. Firstly, we study the density of some functions in $L^1(\mathbb{R}^d)$. By previous results, since $L^1(\mathbb{R}^d)$ is a subset of the space of measurable functions we know that the simple functions must be dense in $L^1(\mathbb{R}^d)$. Similarly, step functions are also dense in $L^1(\mathbb{R}^d)$ since they converge point-wise to f save a set of measure 0, which vanishes under integration. Finally, the continuous functions of compact support are dense in $L^1(\mathbb{R})$. By the observation that step functions are also dense, we observe that it is sufficient to estimate step functions by continuous functions of compact support. In fact, as a step function is simply a finite linear combination of characteristic functions of intervals it suffices to estimate the characteristic function of an interval.

Now, consider the characteristic $\alpha \mathbb{1}_{[a,b]}(x)$ for a real number α. If we are given a sequence of step functions, for the *nth* element in such a sequence, for each characteristic of the aforementioned form attach a function $g(x)$ of a trapezoidal form, becoming steeper and steeper in $n \to \infty$.

Secondly, translation invariance of the integral is heavily useful. Namely, if $f \in L^1(\mathbb{R}^d)$ then

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(x+h)\,\mathrm{d}x, \quad \forall h \in \mathbb{R} \tag{2.22}$$

We shall now show that this is true. The first step is to notice that it is true for simple functions. Indeed, let $\varphi(x) = \sum_{k=1}^N a_k \mathbb{1}_{E_k}(x)$ be a simple function that is also $L^1(\mathbb{R}^d)$. Note that $\varphi(x+h) = \sum_{k=1}^N a_k \mathbb{1}_{E_k}(x+h) = \sum_{k=1}^N a_k \mathbb{1}_{E_k - h}(x)$. However, since the Lebesgue measure μ is *translation invariant* this integral is preserved.

Now, by the density of simple functions, we may take a sequence of simple functions converging to f in $L^1(\mathbb{R}^d)$. So, since these are translation invariant by passing to the limit we may conclude the same for general $L^1(\mathbb{R}^d)$ functions.

2.5 Fubini and Tonelli Theorems

Theorem 2.5.1 (Fubini). *Let f be integrable on $\mathbb{R}^d$, then*

1. *$f^y = f(x,y)$ is integrable for almost all y.*
2. *$\int_{\mathbb{R}^{d_1}} f(x,y)dx$ is integrable for almost all y.*
3. $$\int_{\mathbb{R}^d} f = \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_1}} f(x,y)dydy$$

Proof. Let $\mathcal{F}$ be the space of all functions so that the (3) above hold true. We must show $L^1(\mathbb{R}) \subseteq \mathcal{F}$. We proceed in steps.

1. We now show a finite linear combination is in $\mathcal{F}$. Let $f_1, \ldots f_N$ be a finite collection of elements in $\mathcal{F}$. For each f_j there exists a set A_j so that if $y \notin A_j$ the above holds. Let $A = \bigcup_{j=1}^n A_j$, which also has measure 0. Then, we see that if $y \notin A$ then the above holds for such y. By linearity of the integral, it follows that it holds for all such y.

2. Let us now show the space $\mathcal{F}$ is closed under monotone limits. Indeed, if we select a sequence $(f_n)_n \subset \mathcal{F}$ converging to f almost everywhere, we must show that $f \in \mathcal{F}$. Note that we may assume without loss of generality that each $f_n \geq 0$ by taking $-f_k$ by part (i). Similarly, we may assume the f_n's are increasing by considering $f_1 - f_k$ otherwise.

   With this in mind, let us proceed. Now, each f_k satisfies the criteria for $\mathcal{F}$ save for a set A_k of measure 0. Setting now $A = \bigcup_k A_k$ we see this set has measure 0 and that the criteria are satisfied for each f_k on the complement of this set. Note then that for all $y \notin A$ we have $f_k^y \uparrow f^y$ and hence by the monotone convergence theorem

   $$\int_{\mathbb{R}} f_k^y \to \int_{\mathbb{R}} f_k^y$$

   by monotone convergence. We conclude from this that f_k^y is integrable for almost all y. Similarly, now we have that for almost all y $\displaystyle\int_{\mathbb{R}} f^y(x,y)dx$ is integrable. Another application of monotone convergence (since the f_k's are assumed to be $\uparrow$):

   $$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_k^y dx \right) \to \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f^y dx \right)$$

   whence we have (2) for almost all y for f. Moreover, by assumption

   $$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_k^y(x,y)dx \right) dy = \int_{\mathbb{R}^2} f_k$$

   Again, by monotone convergence we find $\int_{\mathbb{R}^2} f_k \to \int_{\mathbb{R}^2} f$ but also

   $$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f_k^y dx \right) dx \to \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f^y \right) dydx$$

3. We now show any characteristic function is in $\mathcal{F}$. We first consider the case where E is an open cube. That is, $E = Q_1 \times Q_2$. Then,

$$\int_{\mathbb{R}^2} \mathbf{1}_E(x)dx = \mu(E) = \mu(Q_1)\mu(Q_2)$$

For each y the slice $\mathbf{1}_E^y$ is measurable and integrable with

$$\int_{\mathbb{R}} \mathbf{1}_E^y = \begin{cases} \mu(Q_1) & y \in Q_2 \\ 0 & y \notin Q_2 \end{cases}$$

we note this integral is equivalent to the function $\mu(Q_1)\mathbf{1}_{Q_2}$ which is also integrable with

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbf{1}_E dx \right) dy = \int_{\mathbb{R}} \mu(Q_1)\mathbf{1}\mu(Q_2) = \mu(Q_1)\mu(Q_2)$$

finally showing this function is in $\mathcal{F}$.

Now we show the case where f is the characteristic function of a subset of the boundary of a rectangle. In $\mathbb{R}^2$ we must have $\int_{\mathbb{R}^2} \mathbf{f}_E = 0$. For all but finitely many points y, the slice f^y is measurable with

$$\int_{\mathbb{R}} f^y(x)dx = 0$$

and thus is integrable. Now, it follows from this as well that

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y)dx \right) dy = 0$$

as was desired.

We now handle a combination of closed cubes. Assume $E = \bigcup_{k=1}^{N} Q_k$. We may write, for Q_k^0 the interiors of the cubes the characteristic function of E as a finite linear combination of $\mathbf{1}_{Q_k^0}$ and $\mathbf{1}_{\partial Q_k}$ which are each in $\mathcal{F}$. The overall result follows from part (1).

Now let E be an open set with $\mu(E) < \infty$. We may express $E = \bigcup_{n \in \mathbb{N}} Q_n$ for almost disjoint closed cubes Q_n. Now set $f_k = \sum_{n=1}^{N} \mathbf{1}_{Q_n}$ increase to f almost everywhere, whence we conclude by step 2 that f lives in $\mathcal{F}$.

Finally we show this holds for G_δ sets of finite measure. Indeed, if E is such a set then

$$E = \bigcap_n \tilde{O}_n$$

for open sets $\tilde{O}_n$. Since E has finite measure, we have that for some open set $\mathcal{O}_0$ $E \subseteq \mathcal{O}_0$ and hence we see that

$$E = \mathcal{O}_0 \cap \bigcap_n \tilde{O}_n$$

Letting $O_n := \mathcal{O}_0 \cap \bigcap_{k=1}^n \tilde{O}_n$ we have a decreasing sequence of sets of finite measure convering to E. Letting $f_n = f\mathbf{1}_{O_n}$ we have the result by step 2.

4. We show that if E has measure 0 then $\mathbf{1}_E \in \mathcal{F}$. We may take a set $G = \bigcap_n G_n$ of open sets so that $E \subseteq G$ with $\mu(G) = 0$. By the previous step,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_G(x, y) dy dy = \int_{\mathbb{R}} \mathbf{1}_G = 0$$

hence $\int_{\mathbb{R}} \mathbf{1}_G(x, y) = 0$ for almost all y yielding that G^y has measure 0 for almost all y. Then, since $E^y \subseteq G^y$ we see

$$\int_{\mathbb{R}} \mathbf{1}_E(x, y) dx = 0$$

for almost all y. Then,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1} dy dx = 0$$

5. The characteristic function of any set of finite measure lives in $\mathcal{F}$. Indeed, one can find a set G of type G_δ so that $E \subseteq G$ but $\mu(G \setminus E) = 0$. Then, we write $\mathbf{1}_E = \mathbf{1}_G - \mathbf{1}_{G \setminus E}$.

6. Decomposing $f = f^+ - f^-$ and choosing a sequence of simple functions converging to f, we can see that the result holds for L^1 functions in general.

$\square$

Theorem 2.5.2 (Tonelli). *Let f be measurable and non-negative. Then,*

1. f^y is measurable for almost all y.

2. $\int_{\mathbb{R}} f^y dx$ is measurable for almost all y.

3.

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dx \right) dy = \int_{\mathbb{R}^2} f$$

Proof. For such a f, define

$$f_n(x, y) := \begin{cases} f(x, y) & |(x, y)| \leq n, |(x, y)| < n \\ 0 & \text{otherise} \end{cases} \tag{2.23}$$

each f_n is integrable and hence by Fubini we see that for almost all y f_n^y is integrable and

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y) dx dy = \int_{\mathbb{R}^2} f$$

Now, by monotone convergence we have

$$\lim_{n\to\infty} \int_{\mathbb{R}^2} f_n = \int_{\mathbb{R}^2} f \tag{2.24}$$

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n^y = \int_{\mathbb{R}} f^y \tag{2.25}$$

$$\tag{2.26}$$

for all y save a set of measure 0. Moreover, note this secnod line implies the measurability of $\int_{\mathbb{R}} f^y$. Another application of the monotone convergence theorem implies

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f^y(x) dx dy = \lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} f_n^y(x) dx = \lim_{n\to\infty} \int_{\mathbb{R}^n} f_n$$

□

Chapter 3

Differentiation Theorems

3.1 The Hardy-Littlewood Maximal Operator and the Vitali Lemma

Here we shall study a special operator that acts upon the complete space $L^1(\mathbb{R}^d)$. Let us now define this operator:

Definition 10 (Hardy-Littlewood)**.** *Let $f \in L^1(\mathbb{R}^d)$, we define for open balls $B \subseteq \mathbb{R}^d$*

$$f^\star(x) := \sup_{B \ni X} \frac{1}{\mu(B)} \int_B |f(y)| \, \mathrm{d}y, \quad x \in \mathbb{R}^d \tag{3.1}$$

Lemma 3.1.1 (Vitali Covering Lemma)**.** *Let $\{B_1, B_2, \cdots, B_N\}$ be a collection of balls. There exists a disjoint sub-collection of balls $B_{n_1}, \cdots, B_{n_k}$ so that*

$$\mu\left(\bigcup_{j=1}^{N} B_j\right) \leq 3^d \sum_{i=1}^{k} \mu\left(B_{n_k}\right) \tag{3.2}$$

Proof. Let B_{n_1} be the largest ball in the given collection. Delete from the collection all the other balls that have non-empty intersection with B_{n_1}. Dilate this ball by a factor of 3 in radius, all the deleted balls are contained in this expanded ball. Proceeding recursively, we repeat this same process until we run out of balls, which will always terminate as we are considering only finitely many balls. □

Some properties are in order:

Theorem 3.1.2. *The following hold for the function $f^\star(x)$:*

1. *$f^\star$ is measurable in x.*
2. *$f^\star(x) < \infty$ for almost all x.*

3. For $f^\star$ one has for any $\alpha > 0$

$$\mu\left(\left\{x \in \mathbb{R}^d : f^\star(x) > \alpha\right\}\right) \leq \frac{A}{\alpha} \|f\|_{L^1(\mathbb{R}^d)} \tag{3.3}$$

for some constant A.

Proof. Let $a \in \mathbb{R}$, we shall show that $\{f^\star > a\}$ is open. Indeed, if $f^\star(x) > a$ then there must exist a ball $B \ni x$ so that

$$\frac{1}{\mu(B)} \int_B |f(y)| \, \mathrm{d}y > \alpha$$

Then, for any z in this same ball we must then have

$$\frac{1}{\mu(B)} \int_B |f(y)| \, \mathrm{d}y > \alpha$$

and hence have that $B \subseteq \{f > a\}$. Note that the estimate in (3) yields the result in (2), by taking $\alpha \to \infty$, so we need only show (3). Now, let $\alpha > 0$ be given and fix a compact subset K of the set considered. For each $x \in K$ take a ball B_x as above. Therefore, we can construct an open covering $K \subseteq \bigcup_{x \in K} B_x$. By compactness, we may extract a finite subcovering $B_1, B_2, \cdots, B_N$ also containing K in their union. Then,

$$\mu(K) \leq \mu\left(\bigcup_{i=1}^N B_i\right)$$

Now, by the Vitali Covering Lemma we may extract further a disjoint sub-collection $B_{n_1}, B_{n_2}, \cdots, B_{n_k}$ so that $\mu\left(\bigcup_{i=1}^N B_i\right) \leq 3^d \sum_{j=1}^k \mu(B_{n_j})$. Then,

$$\begin{aligned} \mu(K) \leq 3^d \sum_{j=1}^k \mu(B_{n_j}) < \frac{3^d}{\alpha} \sum_{j=1}^k \int_{B_{n_j}} |f(y)| \, \mathrm{d}y &= \frac{3^d}{\alpha} \int_{\bigcup_{j=1}^k B_{n_j}} |f(y)| \, \mathrm{d}y \\ &\leq \frac{3^d}{\alpha} \|f\|_{L^1(\mathbb{R}^d)} \end{aligned}$$

Since this holds for any compact set $K \subseteq \left\{x \in \mathbb{R}^d : f^\star(x) > \alpha\right\}$ we are free to expand K. In the limit, we find the desired result. □

3.2 The Lebesgue Differentiation Theorems

Lemma 3.2.1. *Let f be non-negative, and let $E := \{x : f(x) > \alpha\}$. Then,*

$$\mu(E) \leq \frac{1}{\alpha} \int_E f$$

Proof. This is clear from:

$$\int_E f(x)\,\mathrm{d}x \geq \int_E \alpha\,\mathrm{d}x = \alpha\mu(E)$$

□

Theorem 3.2.2 (Lebesgue Differentiation Theorem)**.** *Suppose $f \in L^1(\mathbb{R}^d)$, then*

$$\lim_{\mu(B)\to 0} \frac{1}{\mu(B)} \int_B f(y)\,\mathrm{d}y = f(x)$$

for almost all x; and the balls B contain x.

Proof. Fix $\alpha > 0$, it is sufficient to show that

$$\mu\left(\left\{\limsup_{\mu(B)\to 0}\left|\frac{1}{\mu(B)}\int_B f(y)\,\mathrm{d}y - f(x)\right| > 2\alpha\right\}\right) = 0$$

denote this set by E. Now, given $\varepsilon > 0$ by their density we may select a continuous function of compact support $g \in L^1(\mathbb{R}^d)$ so that $\|f-g\|_{L^1(\mathbb{R}^d)} < \varepsilon$. Then, we may write

$$\begin{aligned}\frac{1}{\mu(B)}\int_B f(y)\,\mathrm{d}y - f(x) &= \frac{1}{\mu(B)}\int_B (f(y)-g(y))\,\mathrm{d}y - f(x) &(3.4)\\ &+\frac{1}{\mu(B)}\int_B g(y)\,\mathrm{d}y - g(x) + g(x) - f(x) &(3.5)\end{aligned}$$

In the limit superior we find:

$$\limsup_{\mu(B)\to 0}\left|\frac{1}{\mu(B)}\int_B f(y)\,\mathrm{d}y - f(x)\right| \leq (f-g)^\star(x) + |f(x)-g(x)| \tag{3.6}$$

So, we have that $E \subseteq \{(f-g)^\star(x) > \alpha\} \cup \{f-g > \alpha\}$. Thence,

$$\begin{aligned}\mu(E) &\leq \mu\left(\{(f-g)^\star(x) > \alpha\}\right) + \mu\left(\{f-g>\alpha\}\right)\\ &\leq \frac{A}{\alpha}\|f-g\|_{L^1(\mathbb{R}^d)} + \frac{1}{\alpha}\|f-g\|_{L^1(\mathbb{R}^d)}\end{aligned}$$

which can be made arbitrarily small independent of α.

□

Proposition 3.2.3. *If F is increasing and continuous on a subset of $\mathbb{R}$, then F' exists almost everywhere. Moreover, F' is measurable and non-negative with*

$$\int_a^b F'(x)\,\mathrm{d}x \leq F(x) - F(a) \tag{3.7}$$

Proof. Given $n \in \mathbb{N}$ we set

$$G_n(x) := \frac{F(x+1/n) - F(x)}{1/n}$$

as $n \to \infty$ this tends to $G_n(x)$ □

Lemma 3.2.4. *Let F be absolutely continuous on an interval $[a,b]$ and suppose that F' exist almost everywhere on $[a,b]$ and vanishes whenever it exists, save possibly a set of measure 0. Then F is constant on $[a,b]$.*

Proof. Let E be the set of all points in $[a,b]$ where F' exists and is zero, then $\mu(E) = b-a$. Now, fix $\varepsilon > 0$. We construct a Vitali covering $\mathfrak{B}$ as follows: if $x \in E$ then there exists $\delta > 0$ and an interval $(a_x, b_x) \subseteq [a,b]$ with

$$|F(b_x) - F(a_X)| \leq \varepsilon \, |b_x - a_x| \, , \quad b_x - a_x < \delta$$

Then, given $\eta > 0$ by taking a small enough ball about x we can preserve the above. The collection of all such balls, associated to each x, defines a Vitali covering of E $\mathfrak{B}$. By a covering lemma that is well known, given $\delta > 0$ there exists $\{B_j\}_{j=1}^N \subseteq \mathfrak{B}$ so that

$$\mu\left(E - \bigcup_{j=1}^{N} B_j\right) \leq 2\delta$$

since $\mu(E) = \mu(E \cap A) + \mu(E \cap A^c)$ implies that $\mu(E) - \mu\left(E \cap \bigcup_{j=1}^N B_j\right) \leq 2\delta$. That is,

$$\sum_{j=1}^{N} (b_j - a_j) \geq \mu(E) - 2\delta = (b-a) - 2\delta \tag{3.8}$$

However,

$$\sum_{i=1}^{N} |F(b_j) - F(a_j)| \leq \varepsilon(b-a)$$

since the intervals are disjoint, and their sum must be a subset of $[a,b]$. Now, consider the complement of $\bigcup_{j=1}^N (a_j, b_j)$, say $\bigcup_{j=1}^K [\alpha_j, \beta_j]$ which we restrict to $[a,b]$. That is, we take the complement in $[a,b]$. Then, the sum of the measures must add up to $[a,b]$, thus we see that the measure of this finite collection of closed sets is no larger than δ.

Now, we may refine this δ as desired so that

$$\sum_{j=1}^{M} |F(b_j) - F(a_j)| < \varepsilon$$

which we may do by absolute continuity. Then, by the triangle inequality:

$$|F(b) - F(a)| \leq \sum_{j=1}^{N} |F(b_j) - F(a_j)| + \sum_{j=1}^{M} |F(b_j) - F(a_j)| \leq (b - a + 1)\varepsilon$$

which completes the proof. $\square$

Lemma 3.2.5 (Rising Sun Lemma)**.** *Let $G : A \subseteq \mathbb{R} \to \mathbb{R}$ be continuous, and denote by E the collection*

$$E := \{x \in A \mid \exists h_x > 0 \text{ such that } G(x + h_x) > G(x)\} \tag{3.9}$$

If E is non-empty, then $E = \bigcup_n (a_n, b_n)$ for $G(b_n) = G(a_n)$.

Proof. Since E is open by continuity of G, we may write by Theorem (1.1.1) $E = \bigcup_n (a_n, b_n)$ for disjoint open intervals. Since $a_n \notin E$ for any index n, there does not exist suitable $h_x > 0$ and so $G(a_n) \geq G(b_n)$ for all such indices. What we shall show is that $G(a_n) < G(b_n)$ is impossible.

Certainly, suppose that $G(a_n) < G(b_n)$. Then we may find $c \in (a_n, b_n)$ by continuity so that

$$G(c) = \frac{G(a_n) + G(b_n)}{2}$$

Note that $G(b_n) < G(c)$ and so b_n is not a candidate, therefore we may choose the largest such c while maintaining $c < b_n$. Consequently, $c \in E$ so we may choose $d > c$ so that $G(d) > G(c)$. As $b_n \notin E$ we then have $G(x) \leq G(b_n)$ for all $x \geq b_n$. Thus we must have $d < b_n$. Since $G(c) > G(b_n)$ and $b_n > d$ we may find $c' \in (d, b_n)$ so that $G(c') = G(c)$, contradicting our maximal choice. This concludes the proof. $\square$

Theorem 3.2.6 (Fundamental Theorem of Calculus)**.** *Let $F : [a, b] \to \mathbb{R}$ be absolutely continuous. Then F' exists almost everywhere and is integrable. Moreover,*

$$F(x) - F(a) = \int_a^x F'(y)\, \mathrm{d}y \tag{3.10}$$

for all $a \leq x \leq b$. Conversely, if f is integrable on $[a, b]$ then there exists an absolutely continuous function such that $F'(x) = f(x)$ almost everywhere, and the above satisfies this relation.

Proof. Note that if F is absolutely continuous then it is known that it is the difference of two continuous, increasing functions. Then, we have that F' exists and is integrable on $[a, b]$. Let us now define $G(x) = \int_a^x F'(y)\, \mathrm{d}y$. An application of the *Lebesgue* differentiation theorem shows that $G'(x) = F'(x)$ almost everywhere on $[a, b]$ and moreover, we have by the previous lemma that $G(x) = F(x) + C$ on $[a, b]$. Taking $x = a$ yields $C = -F(a)$ which completes the proof.

To see the converse, note that $G(x)$ as defined above is absolutely continuous and the Lebesgue differentiation theorem gives the desired result. $\square$

Chapter 4

Abstract Measure and $L^2(\mathbb{R}^d)$ Space

In this chapter we address in a compact, and elegant form, extensions to the previous results. Recall that we very briefly introduced the concept of a *measure space* in the first Chapter. What we shall do here is refine this concept of a measure on an abstract pair $(X, \mathcal{M})$: for a σ-Algebra $\mathcal{M}$ on X. We begin with the usual definition:

4.1 Abstract Measure Spaces

Definition 11 (Exterior Measure)**.** *Let X be a set, a set function $\mu^\star : 2^X \to [0, \infty]$ is called an **exterior measure** on X provided each of the following hold:*

1. $\mu^\star(\emptyset) = 0$.
2. *If $E_1 \subseteq E_2$ then $\mu^\star(E_1) \leq \mu^\star(E_2)$.*
3. *If $\{E_n\}_{n=1}^\infty$ is a collection of subsets of X then*

$$\mu^\star\left(\bigcup_{n=1}^\infty E_n\right) \leq \sum_{n=1}^\infty \mu^\star(E_n) \tag{4.1}$$

Of course, just as with the Lebesgue measure, for consistency we must restrict ourselves to a class of measurable functions. We shall call a subset of X, E, **Caratheodory** measurable if

$$\mu^\star(A) = \mu^\star(E \cap A) + \mu^\star(E^c \cap A), \quad \text{for all } A \subseteq X \tag{4.2}$$

The collection of all such subsets of X shall be denoted by $\mathcal{M}$ and we shall replace $\mu^\star$ by μ.. The next theorem is absolutely required:

Theorem 4.1.1. *Let X be a set and $\mu^\star$ an exterior measure on X. As defined above, the triple $(X, \mathcal{M}, \mu)$ is a measure space.*

Proof. The majority of the work involves proving that $\mathcal{M}$ is a σ-Algebra. Note that the Caratheodory condition must hold for $\emptyset, X$ and so we already have this (trivial) condition. Now, consider two disjoint sets $E_1, E_2 \in \mathcal{M}$. What we shall now show is that $E_1 \cup E_2 \in \mathcal{M}$. Let $A \subseteq X$ be given and write:

$$\begin{aligned}
\mu^\star(A) &= \mu^\star(E_2 \cap A) + \mu^\star(E_2^c \cap A) \\
&= \mu^\star(E_1 \cap E_2 \cap A) + \mu^\star(E_1^c \cap E_2 \cap A) + \mu^\star(E_1 \cap E_2^c \cap A) + \mu^\star(E_1^c \cap E_2^c \cap A) \\
&= \mu^\star(E_1^c \cap E_2 \cap A) + \mu^\star(E_1 \cap E_2^c \cap A) + \mu^\star(E_1^c \cap E_2^c \cap A) \\
&= \mu^\star(E_1^c \cap E_2 \cap A) + \mu^\star(E_1 \cap E_2^c \cap A) + \mu^\star((E_1 \cup E_2)^c \cap A) \\
&\geq \mu^\star((E_1 \cup E_2) \cap A) + \mu^\star((E_1 \cup E_2)^c \cap A)
\end{aligned}$$

where this last line is due to the fact that $(E_1 \cup E_2) = (E_1^c \cap E_2) \cup (E_1 \cap E_2^c)$ since $E_1 \cap E_2 = \emptyset$. The above inequalities then implies the desired equality and that $E_1 \cup E_2 \in \mathcal{M}$.

We now need to extend this to the countable union $\{E_k\}_{k=1}^\infty$ of disjoint sets in $\mathcal{M}$. For the sake of clarity define

$$G_n := \bigcup_{k=1}^n E_k, \quad G := \bigcup_{k=1}^\infty E_k$$

Now, for each n we have that G_n is the union of finitely many disjoint measurable sets, thus for each $A \subseteq X$:

$$\begin{aligned}
\mu^\star(G_n \cap A) &= \mu^\star(E_n \cap G_n \cap A) + \mu^\star(E_n^c \cap G_n \cap A) \\
&= \mu^\star(E_n \cap A) + \mu^\star(G_{n-1} \cap A) \\
&= \mu^\star(E_n \cap A) + \sum_{k=1}^{n-1} \mu^\star(E_k \cap A) \\
&= \sum_{k=1}^{n} \mu^\star(E_k \cap A)
\end{aligned}$$

where this second last line follows from successive applications of our argument for 2 disjoint, measurable, sets. Now, again we know by this that $G_n, G_n^c \in \mathcal{M}$ and thus

$$\begin{aligned}
\mu^\star(A) = \mu^\star(G_n \cap A) + \mu^\star(G_n^c \cap A) &= \sum_{k=1}^{n} \mu^\star(E_k \cap A) + \mu^\star(G_n^c \cap A) \\
&\geq \sum_{k=1}^{n} \mu^\star(E_k \cap A) + \mu^\star(G^c \cap A)
\end{aligned}$$

since $G_n \subseteq G \implies G_n^c \supseteq G^c$. Passing to the limit in $n \to \infty$ we may conclude that

$$\mu^\star(A) \geq \sum_{k=1}^{\infty} \mu^\star(E_k \cap A) + \mu^\star(G^c \cap A) \geq \mu^\star(G \cap A) + \mu^\star(G^c \cap A)$$

proving that $G \in \mathcal{M}$. Thus, $\mathcal{M}$ is certainly a σ-Algebra on X. Moreover, taking $A = G$ in the above we recover

$$\mu^\star(G) \geq \sum_{k=1}^{\infty} \mu^\star(E_k \cap G) = \sum_{k=1}^{\infty} \mu^\star(E_k) \geq \mu^\star(G)$$

implying that $\mu^\star$ is countably additive on $\mathcal{M}$. This completes the proof. □

We digress briefly to consider a special class of measure spaces: metric spaces endowed with a Borel σ-Algebra and measure. Let (X, d) be a metric space. The Borel σ-Algebra on X is defined to be the smallest $\sigma - Algebra$ containing all the open sets of X, denoted $\mathcal{B}(X)$. More precisely, if $\mathcal{S}$ is any σ-Algebra on X containing the open sets of X, then

$$\mathcal{B}(X) = \bigcap_{\mathcal{S}} \mathcal{S} \tag{4.3}$$

It is not difficult to prove that the intersection (even uncountably many) of σ-Algebras is itself a σ-Algebra, so this is a valid representation. Given two subsets $A, B \subseteq X$ their distance, denoted $d(A, B)$ is simply:

$$d(A, B) := \inf\{d(x, y) : x \in A,\ y \in B\}$$

We shall call a set function $\mu^\star$ on $\mathcal{B}(X)$ a **metric exterior measure** if it satisfies

$$\mu^\star(A \cup B) = \mu^\star(A) + \mu^\star(B), \quad \text{if } d(A, B) = \delta > 0 \tag{4.4}$$

We consequently have the following:

Theorem 4.1.2. *Let (X, d) be a metric space and $\mu^\star$ a metric exterior measure on $\mathcal{B}(X)$. Then, $(X, \mathcal{B}(X), \mu^\star)$ is a measure space.*

Proof. Note that it is sufficient to show that all closed sets satisfy the above. Indeed, recall that the previous theorem guarantees that the collection of all Caratheodory measurable sets is a σ-Algebra. Then, if all closed sets live in this $\mathcal{M}$, we know then that all the open sets live in $\mathcal{M}$. Since $\mathcal{B}(X)$ is the intersection of any such $\mathcal{M}$, we conclude that $\mathcal{B}(X)$ is also a σ-Algebra.

Now let $F \subseteq X$ be a closed set and $A \subseteq X$ arbitrary. We may assume without harm that $\mu^\star(A) < \infty$. Given any $n \in \mathbb{N}$ let us define the auxiliary set

$$A_n := \left\{ x \in A \cap F^c \mid d(x, F) \geq \frac{1}{n} \right\} \tag{4.5}$$

This obviously defines an increasing sequence of sets, i.e $A_n \subseteq A_{n+1}$ for all n. By our very construction of these, we have $d(A_n, A) \geq 1/n$ whence is follows from using that $(F \cap A) \cup A_n \subseteq A$:

$$\mu^\star(A) \geq \mu^\star\left((F \cap A) \cup A_n\right) = \mu^\star(F \cap A) + \mu^\star(A_n) \tag{4.6}$$

Define for $n \in \mathbb{N}$: $B_n := A_{n+1} \cap A_n^c$, and observe

$$d(B_{n+1}, A_n) \geq \frac{1}{n(n+1)}$$

Certainly, if we let $x \in B_{n+1}$ and have $y \in X$ so that $d(x,y) < \frac{1}{n(n+1)}$. Then, if $z \in F$ we have

$$d(y,z) \leq d(y,x) + d(x,z) < \frac{1}{n(n+1)} + \frac{1}{n+1} = \frac{1}{n}$$

and hence $y \notin A_n$. Thus for any positive integer k:

$$\mu^\star(A_{2k+1}) \geq \mu^\star(B_{2k} \cup A_{2k-1}) = \mu^\star(B_{2k}) + \mu^\star(A_{2k-1})$$

by successive applications one can easily see that:

$$\mu^\star(A_{2k+1}) \geq \sum_{j=1}^{k} \mu^\star(B_{2j}), \quad \mu^\star(A_{2k}) \geq \sum_{j=1}^{k} \mu^\star(B_{2j-1}) \tag{4.7}$$

By our hypothesis on $\mu^\star(A) < \infty$, we have that the two series above must be convergent even as $k \to \infty$. Especially,

$$\mu^\star(A_n) \leq \mu^\star(F^c \cap A) \leq \mu^\star(A_n) + \sum_{k=n+1}^{\infty} \mu^\star(B_k)$$

taking $n \to \infty$ we conclude

$$\lim_{n\to\infty} \mu^\star(A_n) = \mu^\star(F^c \cap A)$$

this together with (4.6) proves the desired result. $\square$

4.2 Signed Measures and Decompositions

Given a measurable space $(X, \mathcal{M})$ a set function ν on $\mathcal{M}$ is called a **signed measure** provided it

1. ν takes value in $(-\infty, \infty]$.
2. Countably additive on $\mathcal{M}$ for disjoint sets.

Let us now note that given a signed measure ν on a measurable space we can always construct a set function μ dominating ν so that the triple $(X, \mathcal{M}, \mu)$ is a measure space. Indeed, define the **total variation** of ν:

$$|\nu|(E) = \sup_{\mathcal{M} \supseteq \{E_j\}_j \supseteq E} \sum_j |\nu(E_j)| \tag{4.8}$$

where $\{E_j\}_j$ is a collection of disjoint subsets of $\mathcal{M}$ containing E in their union. The following result is a consequence:

Theorem 4.2.1. *The total variation of ν, $|\nu|$ is a positive measure on $(X, \mathcal{M})$.*

Proof. To prove that $|\nu|$ is a measure we only need to show countable additivity of disjoint sets in $\mathcal{M}$. Let $\{E_j\}_{j=1}^{\infty}$ be a collection of disjoint subsets of $\mathcal{M}$ with $E = \bigcup_{j \in \mathbb{N}} E_j$. We need to prove

$$|\nu|(E) = \sum_{j=1}^{n} |\nu|(E_j) \iff |\nu|(E) \leq \sum_{j=1}^{n} |\nu|(E_j) \wedge |\nu|(E) \geq \sum_{j=1}^{n} |\nu|(E_j) \tag{4.9}$$

For each E_j, take a partition $\{F_{j,k}\}_{k=1}^{\infty}$ so that $E_j = \coprod_{k=1}^{\infty} F_{j,k}$. Then, we note that $\coprod_{j,k=1}^{\infty} F_{j,k}$ is a partitioning of E itself so that

$$\sum_{j,k=1}^{\infty} |\nu(F_{k,j})| = \sum_j \sum_k |\nu(F_{k,j})| \leq |\nu|(E)$$

we now take the supremum over the partitions $\{F_{j,k}\}$ to deduce that

$$\sum_{k=1}^{\infty} |\nu|(E_k) \leq |\nu|(E)$$

For the reverse inequality, let $\{F_k\}_{k=1}^{\infty}$ be a partition of E. Then, for fixed k and varying j we see $\{F_k \cap E_j\}_{j=1}^{\infty}$ is a partition of F_k since the E_j's are disjoint. In this case, we note that

$$\begin{aligned}
\sum_{k=1}^{\infty} |\nu(F_k)| = \sum_{k=1}^{\infty} \left| \sum_{j=1}^{\infty} \nu(F_k \cap E_j) \right| &\leq \sum_k \sum_j |\nu(F_k \cap E_j)| \\
&= \sum_j \sum_k |\nu(F_k \cap E_j)| \\
&\leq \sum_j |\nu|(E_j)
\end{aligned}$$

taking the supremum over $\{F_k\}_k$ yields the result.

□

Given a measure μ and a *signed* measure ν, we say ν is **absolutely continuous** with respect to μ provided $\nu(E) = 0$ whenever $\mu(E) = 0$. We then write $\nu \ll \mu$. We say ν, μ are **mutually singular**, denoted $\mu \perp \nu$, if ν and μ have disjoint support. That is, there exist subsets $A, B \in \mathcal{M}$ with $A \cap B = \emptyset$ and

$$\mu(E \cap A) = \mu(E), \quad \nu(E \cap B) = \nu(E) \tag{4.10}$$

for all $E \in \mathcal{M}$.

This leads us to the *Radon-Nikodym* theorem:

Theorem 4.2.2 (Radon-Nikodym)**.** *Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space and ν a σ-finite measure on the measurable space $(X, \mathcal{M})$. There exist unique signed measure ν_a, ν_s defined on $(X, \mathcal{M})$ so that $\nu_a \ll \mu$ and $\nu_s \perp \mu$ with $\nu = \nu_a + \nu_s$. Moreover,*

$$\nu_a(E) = \int_E f(x)\,\mathrm{d}\mu(x) \tag{4.11}$$

where f is μ-integrable in the extended sense.

4.3 $L^2(\mathbb{R}^d)$ Space

Let us recall that we discussed $L^1(\mathbb{R}^d)$ space in Chapter 2, as the space of all Lebesgue integrable functions. Here, we shall study a similar space: the space of all *square* integrable functions over $\mathbb{R}^d$. A difference, is we shall now admit *complex* valued functions, as we wish to turn $L^2(\mathbb{R}^d)$ into a complex Hilbert space, which shall be defined soon enough.

Definition 12. *Let $f : E \subseteq \mathbb{R}^d \to \mathbb{C}$ be defined on a measurable set E. One may find real valued functions $u, v : E \to \mathbb{R}$ such that $f(x) = u(x) + iv(x)$ by defining $u := \Re f$ and $v := \Im f$. We shall define its Lebesgue integral:*

$$\int_{\mathbb{R}^d} f(x)\,\mathrm{d}x := \int_{\mathbb{R}^d} u(x)\,\mathrm{d}x + i\int_{\mathbb{R}^d} v(x)\,\mathrm{d}x \tag{4.12}$$

and call f integrable if and only if u, v are integrable in the real sense.

Now, we consider the primitive set

$$\mathcal{L} := \left\{ f : \mathbb{R}^d \to \mathbb{C} : \int_{\mathbb{R}^d} |f(x)|^2 \,\mathrm{d}x < \infty \right\}$$

We define an equivalence relation $\sim$ on $\mathcal{L}$. We shall say $f \sim g$ if and only if $f = g$ almost everywhere. This is a well defined equivalence relation. Certainly, $f \sim f$ for all $f \in \mathcal{L}$. Obviously, $f \sim g \iff g \sim f$. If $f \sim g$ and $g \sim h$ for $f, g, h \in \mathcal{L}$ then we must have $f \sim h$. Indeed, there exist sets $A, B \subset \mathbb{R}^d$ with $\mu(A) = 0 = \mu(B)$ so that $f(x) = g(x)$

for all $x \notin A$ and $g(x) = h(x)$ whenever $x \notin B$. Define $C := A \cup B$, also a null set. For any $x \notin C$ we have $f(x) = g(x) = h(x)$ and hence $f \sim h$.

Now, we define $L^2(\mathbb{R}^d)$ to be the set of all equivalence classes of $\mathcal{L}$ under the relation $\sim$ where we write $f = g$ whenever $f \sim g$. We shall view this as a vector space over the complex numbers by endowing this structure with the usual vector addition and scalar multiplication over $\mathbb{C}$. We also define an *inner product* on $L^2(\mathbb{R}^d)$ as follows:

$$L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \ni (f, g) \mapsto \langle f, g \rangle := \int_{\mathbb{R}^d} f(x)\overline{g(x)}\,\mathrm{d}x \tag{4.13}$$

It is easy to see that for any $f, g \in L^2(\mathbb{R}^d)$ one has $\langle f, g \rangle = \overline{\langle g, f \rangle}$. This induces the L^2 norm, $\|\cdot\|_2$, on $L^2(\mathbb{R}^d)$ given by:

$$\|\cdot\|_2 := \sqrt{\langle \cdot \rangle} = \left(\int_{\mathbb{R}^d} |f(x)|^2 \,\mathrm{d}x \right)^{1/2}$$

Now we define a complex inner product:

Definition 13. *Let $(V, +, \cdot)$ be a vector space over $\mathbb{C}$. A binary operator $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ is called a* ***complex inner product*** *if*

1. *$\langle x, y \rangle = \overline{\langle g, f \rangle}$ for all $x, y \in V$.*
2. *$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ for all $x, y \in V$ and $\lambda \in \mathbb{C}$.*
3. *$\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ for all $x, y, z \in V$.*
4. *$\langle x, x \rangle \geq 0$ for all $x \in V$ and is equal to 0 if and only if $x = 0_V$.*

We may now define a Hilbert space:

Definition 14 (Hilbert Space)**.** *Let $\mathcal{H}$ be a vector space over $\mathbb{C}$ with operations $(+ . \cdot)$ and let $\langle \cdot, \cdot \rangle$ be a complex inner product on $\mathcal{H}$. Denote by $\|\cdot\|$ the norm on $\mathcal{H}$ induced by the inner product. If we set $d(x, y) := \|x - y\|$ for $x, y \in \mathcal{H}$, and the metric space $(\mathcal{H}, d)$ is complete and separable, then $\mathcal{H}$ is called a* ***Hilbert space****.*

We indeed have some work to do. Namely, we must show that $L^2(\mathbb{R}^d)$ is a Hilbert space. First we show $L^2(\mathbb{R}^d)$ is a complex vector space. Certainly, if $f \in L^2(\mathbb{R}^d)$ we have that $\|f\|_2 < \infty$. Therefore, for any $\lambda \in \mathbb{C}$ one easily sees that $\||\lambda f|\|_2 < \infty$ as well.

If $f, g \in L^2(\mathbb{R}^d)$, then we have $\|f\|_2 < \infty$ and $\|g\|_2 < \infty$. Note that

$$|f + g|\,(x) \leq 2 \max(|f(x)|\,, |g(x)|)$$

Now, consider $|f + g|^2$. Then by the above we have

$$|f + g|^2 \leq 4 \left(|f|^2 + |g|^2 \right)$$

Whence,

$$0 \leq \int_{\mathbb{R}^d} |f(x) + g(x)|^2 \, \mathrm{d}x \leq 4 \left(\int_{\mathbb{R}^d} |f(x)|^2 \, \mathrm{d}x + \int_{\mathbb{R}^d} |g(x)|^2 \, \mathrm{d}x \right)$$

or, $\|f+g\|_2^2 \leq 4\|f\|_2^2 + 4\|g\|_2^2 < \infty$. It is obvious now that $(f+g) \in L^2(\mathbb{R}^d)$. Hence, $L^2(\mathbb{R}^d)$ is a vector space over the complex numbers. We now turn to the inner product (yet to be established). For this to be well defined, we must have $f(x)\overline{g(x)}$ integrable whenever f, g are.

Indeed, it is easy to show that for non-negative $x, y \in \mathbb{R}$ one has $2xy \leq x^2 + y^2$, as is easily proven by considering $(x-y)^2 \geq 0$. Hence,

$$\int_{\mathbb{R}^d} \left| f(x)\overline{g(x)} \right| \, \mathrm{d}x \leq \frac{1}{2} \int_{\mathbb{R}^d} |f(x)|^2 \, \mathrm{d}x + \frac{1}{2} \int_{\mathbb{R}^d} |g(x)|^2 \, \mathrm{d}x < \infty$$

since both $\|f\|_2$ and $\|g\|_2$ were assumed to be finite.

Now, if $\|f\|_2 = \|g\|_2 = 0$ then clearly $|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2$. If $\|f\|_2 = \|g\|_2 = 1$ then we may write

$$|\langle f, g\rangle| \leq \int_{\mathbb{R}^d} \left| f(x)\overline{g(x)} \right| \, \mathrm{d}x \leq \frac{1}{2} \left(\|f\|_2 + \|g\|_2 \right)$$

by the above estimate, again yielding $|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2$. Now, if both norms are non-zero we define

$$f' := \frac{f}{\|f\|_2}, \quad g' := \frac{g}{\|g\|_2}$$

so that $\|f'\| = 1$ and $\|g'\| = 1$. By our previous case, we see $|\langle f', g'\rangle| \leq 1$. Now, multiplying through by $\|f\|_2 \cdot \|g\|_2$ yields the Cauchy-Schwarz inequality. Finally, to establish the triangle inequality, write by Cauchy-Schwarz:

$$\begin{aligned} \left| \|f+g\|_2^2 \right| \leq \langle f+g, f+g\rangle &= \|f\|_2^2 + |\langle f, g\rangle| + |\langle g, f\rangle| + \|g\|_2^2 \\ &= \|f\|_2^2 + 2\,|\langle g, f\rangle| + \|g\|_2^2 \\ &\leq \|f\|_2^2 + 2\,\|f\|_2^2\,\|g\|_2^2 + \|g\|_2^2 \\ &= \left(\|f\|_2^2 + \|g\|_2^2 \right)^2 \end{aligned}$$

To show that $L^2(\mathbb{R}^d)$ is a Hilbert space under these structures, it remains to show that the induced metric space is both complete and separable.

Theorem 4.3.1. *The space* $\left(L^2(\mathbb{R}^d), \|\cdot\|_2 \right)$ *is complete.*

Proof. Let $(f_n)_n \subset L^2(\mathbb{R}^d)$ be a Cauchy sequence. There then exists a subsequence (f_{n_k}) so that for all indices k one has $\|f_{n_{k+1}} - f_{n_k}\| < 2^{-k}$. Fix such a subsequence, and define $f : \mathbb{R}^d \to \mathbb{C}$:

$$f(x) := f_{n_1}(x) + \sum_{k=1}^{\infty} \left(f_{n_{k+1}}(x) - f_{n_k}(x)\right) \tag{4.14}$$

Note that the k^{th} partial sum of the expression on the right-hand side is precisely the point f_{n_k}. We claim that $f \in L^2(\mathbb{R}^d)$. This is consequence of our estimates above, indeed for any $N \in \mathbb{N}$:

$$\|f\|_2 \leq \|f_{n_1}\|_2 + \sum_{k=1}^{N} \|f_{n_{k+1}} - f_{n_k}\|_2 \leq \|f_{n_1}\|_2 + \sum_{k=1}^{N} 2^{-k}$$

letting $N \to \infty$ we recover $\|f\|_2 \leq \|f_{n_1}\|_2 + \sum_{k=1}^{\infty} 2^{-k} < \infty$ by hypothesis. Moreover, it follows that $|f|^2$ is integrable, and hence $|f| < \infty$ almost everywhere. More precisely, the series on the right-hand side of (4.14) converges for almost all $x \in \mathbb{R}^d$ whereby one has that

$$\lim_{k\to\infty} f_{n_k}(x) = f(x)$$

for almost all x, since the f_{n_k} are the partial sums of the series. Now, to establish convergence in this space:

$$\begin{aligned}\|f_{n_k} - f\|_2^2 = \int_{\mathbb{R}^d} \lim_{\ell\to\infty} |f_{n_k}(x) - f_{n_\ell}(x)| &\leq \liminf_{\ell\to\infty} \int_{\mathbb{R}^d} |f_{n_k}(x) - f_{n_\ell}(x)| \\ &= \liminf_{\ell\to\infty} \|f_{n_k} - f_{n_\ell}\|_2^2\end{aligned}$$

which can be made arbitrarily small by taking k large enough, since this sequence is Cauchy. Ergo, given $\varepsilon > 0$ we may find N large enough so that for all $k \geq N$ we have $\|f_{n_k} - f\|_2^2 \leq \varepsilon^2$, i.e. $\|f_{n_k} - f\|_2 \leq \varepsilon$, which completes the proof of the theorem. □

A staggering result is the following:

Theorem 4.3.2. *The metric space $L^2(\mathbb{R}^d, \|\cdot\|_2)$ is separable. In other words, there exists a countable subset of $L^2(\mathbb{R}^d)$ whose linear combinations are dense in $L^2(\mathbb{R}^d)$.*

Proof. Let $f \in L^2(\mathbb{R}^d)$ be given. Define a sequence of functions for $n \in \mathbb{N}$ given by

$$g_n(x) := \begin{cases} f(x) & |x| \leq n \ \& \ |f(x)| \leq n \\ 0 & \text{else} \end{cases}$$

Now, since $f \in L^2(\mathbb{R}^d)$ it follows that f is finite almost everywhere, and therefore we have that $g_n(x) \to f(x)$ for almost all x as $n \to \infty$. Moreover, the difference $|g_n - f|^2$ is bounded above by $2|f|^2$, which is integrable since $f \in L^2(\mathbb{R}^d)$. Therefore, the *Dominated Convergence Theorem* yields that

$$\lim_{n\to\infty} \|g_n - f\|_2 = 0$$

Hence for any $\varepsilon > 0$ there is N so that $\|g_N - f\| < \varepsilon$. Therefore, it suffices to approximate these bounded functions supported on compact cubes.

Consider the family of functions $\{r\mathbb{1}_Q\}$ where $r \in \mathbb{Q}$ and $Q := (r_1, r_2, \cdots, r_d)$ is a cube with rational coordinates. Clearly, this family is countable since $\mathbb{Q}^d$ is countable. Fix $\varepsilon > 0$. By their density in $L^1(\mathbb{R}^d)$ there exists a step function, say, φ satisfying

$$|\varphi(x)| \leq N, \quad \|\varphi - f\|_{L^1(\mathbb{R}^d)} < \frac{\varepsilon^2}{2N} \tag{4.15}$$

In this case,

$$\|\varphi - f\|_2^2 = \int_{\mathbb{R}^d} |\varphi() - f(x)|^2 \, \mathrm{d}x = \int_{\mathbb{R}^d} |\varphi(x) - f(x)| \, \overline{|\varphi(x) - f(x)|} \, \mathrm{d}x \tag{4.16}$$

$$\leq 2N \int_{\mathbb{R}^d} |\varphi(x) - f(x)| \, \mathrm{d}x \tag{4.17}$$

$$\leq 2N \cdot \frac{\varepsilon^2}{2N} = \varepsilon^2 \tag{4.18}$$

Whence $\|\varphi - f\|_2 < \varepsilon$ as was required. $\square$

Bibliography

[1] E. Stein, R. Shakarchi. *Measure Theory, Integration and Hilbert Spaces*, Princeton Lectures in Analysis, 3rd edition, 2007.

[2] E. Stein, R. Shakarchi. *Complex Analysis*, Princeton Lectures in Analysis, 3rd edition, 2007.

[3] Partial Differential Equations, Lawrence C. Evans. volume 19, Graduate Studies in Mathematics.

[4] Partial Differential Equations, Jurgen Jost. Graduate Texts in Mathematics. ISBN: 978-1-4614-4809-9; Springer.

[5] I.G. Petrovskii, Partial Differntial Equations. Moscow: Saunders Company, 1967.

[6] B. Gidas, Wei-Ming Ni, I. Nirenberg. `Symmetry of Positive Solutions of Nonlinear Elliptic Equations in` $\mathcal{R}^n$.