2010

The Riesz representation theorem for linear functionals

Thomas Daniel Schellhous

The Riesz Representation Theorem for Linear Functionals

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Mathematics

by

Thomas Daniel Schellhous

September 2010

THE RIESZ REPRESENTATION THEOREM FOR LINEAR FUNCTIONALS

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Thomas Daniel Schellhous

September 2010

Approved by:

[redacted]

Dr. Yuichiro Kakihara, Committee Chair

8/5/2010

Date

[redacted]

Dr. Corey Dunn, Committee Member

[redacted]

Dr. John Sarli, Committee Member

[redacted]

Dr. Peter Williams, Chair,
Department of Mathematics

[redacted]

Dr. Joseph Chavez
Graduate Coordinator,
Department of Mathematics

Abstract

This project is an expository survey of the Riesz representation theorem for linear functionals, which states that on locally compact Hausdorff spaces, certain linear functionals can be represented by a unique regular Borel measure. This representation is realized in the sense that applying the linear functional to any function in a specific class of functions is identical to integrating the function with respect to the Borel measure. Preliminary material and applications in the areas of measure theory, integration theory, topology, and functional analysis are discussed and thoroughly investigated prior to the statement and proof of the theorem, which is presented in its entirety.

Acknowledgements

First and foremost, I would like to thank Dr. Kakihara for all of his patient guidance. I know that, at times, I really put that patience to the test, but he was always available and willing to help me through any kind of rough patch I might have been going through. I really admire his command of the subject material, and I know that I could never have completed this project without him. My thanks also go to my advisors, Dr. Dunn and Dr. Sarli, for all of their advice and suggestions. Without Dr. Dunn's mathematics typesetting help and experience, this thesis and especially my presentation would not have turned out as well as they did.

I would also like to thank my family for putting up with my late-night study marathons, my all-nighters spent frantically typing away at my thesis, and my resolute, and ultimately incorrect, assumption that math is as interesting to everyone else as it is to me, and that you really wanted to hear all about the latest proof or theorem that I was working on.

Finally, I would like to thank all of my friends for all the help and support they've given me, and especially my best friend Shahla Sadjadi, who inspires me to be better in all aspects of my life. Without her love and support, I would most likely not have finished this project for another ten years, if at all!

Table of Contents

Abstract iii

Acknowledgements iv

1 Introduction 1
1.1 Summary 1
1.2 Motivation 2

2 Measures 3
2.1 σ-algebras 3
2.2 Measures 4
2.3 Properties of Lebesgue Measure 9

3 Integration 14
3.1 Measurable Functions 14
3.2 Integration 19
3.3 Integral Limit Theorems 26

4 The Riesz Representation Theorem for Linear Functionals 33
4.1 Preliminary Material 33
4.2 The Riesz Representation Theorem 38
4.3 Proof of the Riesz Representation Theorem 41

5 Conclusion 55

Bibliography 57

Chapter 1

Introduction

1.1 Summary

The following work investigates the Riesz representation theorem for linear functionals in relation to locally compact Hausdorff spaces. Two other theorems that are commonly called the "Riesz representation theorem" are the theorem for finite-dimensional inner product spaces and the theorem for Hilbert spaces [BN00], and studying these interesting topics helps us to not only gain a better understanding of how linear functionals interact with the vector spaces over which they are defined, but also to start to see faint threads that hint at a deep connection between the various fields of modern mathematics. The representation theorem that we will study establishes an important and remarkable relationship between linear functionals and integration. Informally, the theorem states that given a positive linear functional over the vector space $\mathscr{K}(X)$ of real-valued continuous functions with compact support defined on a locally compact Hausdorff space X, there exists a unique regular Borel measure so that applying the linear functional to any function in $\mathscr{K}(X)$ is identical to integrating that function over the space X with respect to the measure.

Before we get to the theorem, we first introduce and study some preliminary concepts, starting with σ-algebras and measures, and prove some necessary results that will be used throughout this work. Once we have these tools at our disposal, we will then construct the Lebesgue integral in a three-step process. After some fundamental properties of the integral are presented, we move on to some extremely powerful limit theorems that describe situations in which integration can commute with the limit process as applied to convergent sequences of functions. We will then briefly investigate some topological notions, including local compactness and Hausdorff spaces, both of which are essential parts of the main theorem of this project. Once this introductory material is completed, we will be in a position to state and prove the Riesz representation theorem, which is our goal.

1.2 Motivation

One motivating factor for the study of measure theory begins with the desire to generalize the notion of "length" to sets for which the standard definition of Euclidean length does not apply. One example of this would be the rationals, $\mathbb{Q}$. Since $\mathbb{Q}$ contains no intervals, it is not obvious how one would go about discussing length in regards to this set. If we wanted to generalize the notion of length to arbitrary subsets of $\mathbb{R}$, we would want a few things. First, we would want the "size" of a set to be greater than or equal to that of any of its subsets, and this size should be non-negative. Another property that seems logical is that if we have a collection of disjoint sets, the size of the union of the sets should be equal to the sum of the sizes of each individual set. It would also be very nice if this new idea of "size" agreed with the Euclidean definition of length on sets to which it applies.

Unfortunately, there does not exist a function that assigns to each subset of $\mathbb{R}$ a size that meets all of the above requirements [Gor94]. If we want it to match the Euclidean distance and add up properly, there is no way to extend this idea to *every* subset. Alternatively, if we want it to add up properly and apply to every set, then there is no way to make it agree with Euclidean length. However, if we restrict our attention to a specific collection of subsets of $\mathbb{R}$, the Borel σ-algebra, then there is such a function, called the Lesbesgue measure. We can also generalize these concepts to arbitrary sets rather than just $\mathbb{R}$. Once we do this, the ideas of integration and convergence follow naturally as we re-examine the well-established ideas of basic analysis through the lens of measure theory.

Chapter 2

Measures

In this chapter we will define some of the fundamental concepts we will be using for this project, such as σ-algebras and measure spaces. We will also present some important theorems that are used frequently as tools when dealing with measure spaces, and construct a particularly useful measure known as the Lebesgue measure. We will conclude this chapter by examining some interesting properties and applications of the Lebesgue measure.

2.1 σ-algebras

Definition 2.1.1. Let X be an arbitrary set. A collection $\mathscr{A}$ of subsets of X is called a *σ-algebra* on X if the following conditions hold:

1. $X \in \mathscr{A}$,
2. for every $A \in \mathscr{A}$, we have that $A^c \in \mathscr{A}$, where $A^c = X \setminus A$,
3. for every countable collection $\{A_i\}_{i=1}^{\infty}$ of subsets of $\mathscr{A}$, we have that $\bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$,
4. for every countable collection $\{A_i\}_{i=1}^{\infty}$ of subsets of $\mathscr{A}$, we have that $\bigcap_{i=1}^{\infty} A_i \in \mathscr{A}$.

It can easily be shown that conditions 2 and 3 together imply condition 4, and that conditions 2 and 4 together imply condition 3, so that we really only need three conditions in our definition. Replacing countable collections with finite collections in conditions 3 and 4 changes the definition to that of an *algebra* of sets. Note that it is a direct result of the definition of a σ-algebra that if A and B are elements of a σ-algebra

$\mathscr{A}$, then $A \setminus B \in \mathscr{A}$ as well, since $A \setminus B = A \cap B^c$. The following theorem leads to a useful way to find a σ-algebra containing a desired collection of subsets. The simple proof follows directly from the definition of a σ-algebra.

Theorem 2.1.2. *Let X be a set, and let $\mathscr{C}$ be an arbitrary non-empty collection of σ-algebras on X. If $\mathscr{A} = \bigcap_{C \in \mathscr{C}} C$, then $\mathscr{A}$ is a σ-algebra on X.*

A direct corollary of this theorem is that given any collection $\mathscr{C}$ of subsets of X, there is a smallest σ-algebra on X that includes $\mathscr{C}$, which is the intersection of every σ-algebra that contains $\mathscr{C}$. Note that at the very least, the power set on X is a σ-algebra that includes $\mathscr{C}$, so the intersection is non-empty. This intersection is called the *σ-algebra generated by $\mathscr{C}$*. A particularly useful and important σ-algebra is our next definition.

Definition 2.1.3. The *Borel σ-algebra* on $\mathbb{R}$ is the σ-algebra on $\mathbb{R}$ generated by the collection of all open subsets of $\mathbb{R}$. We denote the Borel σ-algebra by $\mathscr{B}(\mathbb{R})$.

This particular σ-algebra is useful because it contains every open subset of $\mathbb{R}$, every closed subset of $\mathbb{R}$, every compact subset of $\mathbb{R}$, and indeed "virtually every subset of $\mathbb{R}$ that is of interest in analysis" [Coh97]. The Borel σ-algebra is also generated by the collection of all closed subsets of $\mathbb{R}$, which is easily verified using the properties of σ-algebras and the fact that the complement of every open subset of $\mathbb{R}$ is closed in $\mathbb{R}$. The elements of $\mathscr{B}(\mathbb{R})$ are called *Borel subsets* or *Borel-measurable*. In general, elements of a σ-algebra are referred to as *measurable sets*, for reasons which will be clear in the next section.

2.2 Measures

Definition 2.2.1. Let X be an arbitrary set, and let $\mathscr{A}$ be a σ-algebra on X. A *measure* is a function μ which satisfies the following conditions:

1. $\mu : \mathscr{A} \to [0, \infty]$,
2. $\mu(\emptyset) = 0$, and
3. μ is countably additive, so that if $\{A_i\}_{i=1}^{\infty}$ is a countable collection of disjoint sets in $\mathscr{A}$, then
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Now we present some terminology. If X is a set, if $\mathscr{A}$ is a σ-algebra on X, and if μ is a measure on $\mathscr{A}$, then we call the pair $(X, \mathscr{A})$ a *measurable space*, and we call the triple $(X, \mathscr{A}, \mu)$ a *measure space*. A measure μ is called *finite* if $\mu(X) < \infty$, and is called *σ-finite* if X is the union of a countable collection of sets $\{A_i\}$, where $\mu(A_i) < \infty$ for each $i \geq 1$. In these cases we also call the measure space $(X, \mathscr{A}, \mu)$ finite or σ-finite. The following useful property can be helpful when trying to prove things about measures.

Theorem 2.2.2. *Let $(X, \mathscr{A}, \mu)$ be a measure space. If $\{A_k\}$ is an arbitrary sequence of sets that belong to $\mathscr{A}$, then*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Proof. We will use the sequence of sets $\{A_k\}$ to construct a related sequence of disjoint sets as follows. Define $B_1 = A_1$. Now let $B_k = A_k \setminus \left(\bigcup_{i=1}^{k-1} A_i\right)$ for $k > 1$. Since $\mathscr{A}$ is closed under countable unions and set subtraction, $B_k \in \mathscr{A}$ for all $k \geq 1$. The sets B_k are clearly disjoint. Also, since $B_k \subseteq A_k$ for every $k \geq 1$, we have that $\mu(B_k) \leq \mu(A_k)$.

Now, if $x \in \bigcup_{k=1}^{\infty} B_k$, then $x \in B_{k_0}$ for some $k_0 \geq 1$. So $x \in A_{k_0}$ since each B_k is contained in the corresponding A_k. Thus $x \in \bigcup_{k=1}^{\infty} A_k$. If $x \in \bigcup_{k=1}^{\infty} A_k$, then there exists a $k_0 \geq 1$ so that $x \in A_{k_0}$ and $x \notin A_j$ for all $j < k_0$. Thus, $x \in A_{k_0} \setminus \left(\bigcup_{i=1}^{k_0-1} A_i\right) = B_{k_0}$, so $x \in \bigcup_{k=1}^{\infty} B_k$. Therefore, $\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k$ by double containment. It now follows that

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \mu\left(\bigcup_{k=1}^{\infty} B_k\right) = \sum_{k=1}^{\infty} \mu(B_k) \leq \sum_{k=1}^{\infty} \mu(A_k). \qquad \square$$

As mentioned earlier, there is a measure that can be defined on $\mathscr{B}(\mathbb{R})$ that assigns to each subinterval of $\mathbb{R}$ its length. Before we can discuss this, we need a few more preliminary definitions.

Definition 2.2.3. Let X be a set, and let $\mathscr{P}(X)$ be the power set on X (the collection of all subsets of X). An *outer measure* on X is a function $\mu^* : \mathscr{P}(X) \to [0, \infty]$ which satisfies the following conditions:

1. $\mu^*(\emptyset) = 0$,

2. μ^* is monotonic, so that if $A \subseteq B \subseteq X$, then $\mu^*(A) \leq \mu^*(B)$, and

3. μ^* is countably subadditive, so that if $\{A_n\}_{n=1}^{\infty}$ is a countable collection of subsets of X, then
$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

As is evident from the definitions, a measure is an outer measure if and only if the domain of the measure is the power set of X. On the other hand, an outer measure is not necessarily a measure, since countable subadditivity does not necessarily imply countable additivity. The following outer measure on $\mathbb{R}$ is perhaps the most useful one.

Definition 2.2.4. For each subset A of $\mathbb{R}$, let $\mathscr{C}_A$ be the set of all infinite sequences $\{(a_i, b_i)\}$ of bounded open subintervals of $\mathbb{R}$ such that $A \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)$. Now define $\lambda^* : \mathscr{P}(\mathbb{R}) \to [0, \infty]$ by

$$\lambda^*(A) = \inf\left\{\sum_i (b_i - a_i) : \{(a_i, b_i)\} \in \mathscr{C}_A\right\}.$$

λ^* is called the *Lesbesgue outer measure* on $\mathbb{R}$.

The useful thing about Lesbesgue outer measure is that it assigns to each subinterval of $\mathbb{R}$ its length. It is not a measure, however, as it is not countably additive. To solve this problem, we turn to the following definition.

Definition 2.2.5. Let X be a set, and let μ^* be an outer measure on X. A subset $B \subseteq X$ is called *μ^*-measurable* if

$$\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$$

holds for every $A \subseteq X$. A *Lesbesgue measurable* subset of $\mathbb{R}$ is one that is λ^*-measurable.

So a μ^*-measurable set is one that "breaks up" every subset of X in such a way that the measures of each piece add up properly. The fact that these sets are called measurable is not a coincidence, since an outer measure naturally induces a specific σ-algebra over which the outer measure becomes a measure. The following theorem describes this process.

Theorem 2.2.6. *Let X be a set, let μ^* be an outer measure on X, and let $\mathscr{A}_{\mu^*}$ be the collection of all μ^*-measurable subsets of X. Then*

1. $\mathscr{A}_{\mu^*}$ *is a* σ*-algebra, and*

2. *the restriction of* μ^* *to* $\mathscr{A}_{\mu^*}$ *is a measure on* $\mathscr{A}_{\mu^*}$.

We now can define the Lesbesgue measure λ on $\mathbb{R}$ to be the restriction of λ^* to $\mathscr{A}_{\lambda^*}$. This σ-algebra is very large in the sense that it contains most subsets of $\mathbb{R}$ that are mathematically useful. We will prove just how large $\mathscr{A}_{\lambda^*}$ is next, which will require the following lemma.

Lemma 2.2.7. *The Borel* σ*-algebra* $\mathscr{B}(\mathbb{R})$ *on* $\mathbb{R}$ *(which we defined to be generated by the collection of all open subsets of* $\mathbb{R}$*) is also generated by the collection of all subintervals of* $\mathbb{R}$ *of the form* $(-\infty, b]$.

Proof. Let $\mathscr{B}_1$ be the σ-algebra generated by the collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$. First, any interval of the form $(-\infty, b]$ can be written as $(b, \infty)^c \in \mathscr{B}(\mathbb{R})$, so $\mathscr{B}_1 \subseteq \mathscr{B}(\mathbb{R})$.

Now let (a, b) be an open interval in $\mathbb{R}$. If $\{c_i\}$ is a sequence of real numbers less than b so that $\lim_{i\to\infty} c_i = b$, then we can express (a, b) as

$$(-\infty, a]^c \bigcap \left(\bigcup_{i=1}^{\infty} (-\infty, c_i] \right) \in \mathscr{B}_1.$$

So every open subinterval of $\mathbb{R}$ is an element of $\mathscr{B}_1$, and since every open subset of $\mathbb{R}$ is the union of a sequence of open subintervals, every open subset of $\mathbb{R}$ is an element of $\mathscr{B}_1$, so $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{B}_1$.

Thus, $\mathscr{B}(\mathbb{R}) = \mathscr{B}_1$. □

Theorem 2.2.8. *Every Borel subset of* $\mathbb{R}$ *is Lebesgue measurable.*

Proof. We will first prove that any interval of the form $(-\infty, b]$ is Lebesgue measurable. Let B be an interval of the form $(-\infty, b]$ and let $A \subseteq \mathbb{R}$. Since λ^* is countably subadditive and $A \subseteq (A \cap B) \cup (A \cap B^c)$, it is automatically true that

$$\lambda^*(A) \leq \lambda^*(A \cap B) + \lambda^*(A \cap B^c).$$

So we only need to prove that

$$\lambda^*(A) \geq \lambda^*(A \cap B) + \lambda^*(A \cap B^c),$$

which is necessarily true if $\lambda^*(A) = \infty$.

So, assume that $\lambda^*(A) < \infty$. Let $\epsilon > 0$, and let $\{(a_i, b_i)\}$ be a sequence of open subintervals of $\mathbb{R}$ that covers A and is such that

$$\sum_{i=1}^{\infty}(b_i - a_i) < \lambda^*(A) + \epsilon.$$

For each $i \geq 1$, the sets $(a_i, b_i) \cap B$ and $(a_i, b_i) \cap B^c$ must be disjoint intervals, which may be empty. Thus, we can cover these sets with open intervals by choosing, for each i, intervals (c_i, d_i) and (e_i, f_i) so that

$$((a_i, b_i) \cap B) \subseteq (c_i, d_i),$$

$$((a_i, b_i) \cap B^c) \subseteq (e_i, f_i), \text{ and}$$

$$(d_i - c_i) + (f_i - e_i) \leq b_i - a_i + \frac{\epsilon}{2^i}.$$

Now, the sequence $\{(c_i, d_i)\}$ covers the set $A \cap B$, and the sequence $\{(e_i, f_i)\}$ covers the set $A \cap B^c$. Therefore, the definition of λ^* leads to the following inequalities:

$$\lambda^*(A \cap B) \leq \sum_{i=1}^{\infty}(d_i - c_i),$$

$$\lambda^*(A \cap B^c) \leq \sum_{i=1}^{\infty}(f_i - e_i).$$

But by above,

$$\begin{aligned}
\sum_{i=1}^{\infty}(d_i - c_i) + \sum_{i=1}^{\infty}(f_i - e_i) &= \sum_{i=1}^{\infty}(d_i - c_i + f_i - e_i) \\
&\leq \sum_{i=1}^{\infty}\left(b_i - a_i + \frac{\epsilon}{2^i}\right) \\
&= \sum_{i=1}^{\infty}(b_i - a_i) + \epsilon.
\end{aligned}$$

Therefore, we have that

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \sum_{i=1}^{\infty}(b_i - a_i) + \epsilon < \lambda^*(A) + 2\epsilon,$$

by our choice of the sequence $\{(a_i, b_i)\}$. Since ϵ was arbitrary, we must have that

$$\lambda^*(A) \geq \lambda^*(A \cap B) + \lambda^*(A \cap B^c),$$

so B is Lebesgue measurable.

Therefore, the collection $\mathscr{A}_{\lambda^*}$ of Lebesgue measurable sets is a σ-algebra on $\mathbb{R}$ containing every subinterval of $\mathbb{R}$ of the form $(-\infty, b]$. But by Lemma 2.2.7, $\mathscr{B}(\mathbb{R})$ is the smallest σ-algebra containing all subintervals of this form, so we must have that $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{A}_{\lambda^*}$. Thus, every Borel subset of $\mathbb{R}$ is Lebesgue measurable. $\square$

2.3 Properties of Lebesgue Measure

Earlier, we discussed the desire to generalize the notion of "size" or "length" to more arbitrary subsets of $\mathbb{R}$. Lebesgue measure λ is one way to try to solve this problem. In fact, Lebesgue measure is the only measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ that assigns to every interval its length [Coh97]. Furthermore, in a certain sense, Lebesgue measure is the "natural" measure on $\mathscr{B}(\mathbb{R})$, since every positive translation-invariant Borel measure on $\mathbb{R}$ that is finite on compact sets is a constant multiple of λ [Rud87]. We will now present a few interesting facts about Lebesgue measure.

Theorem 2.3.1. *Let A be a Lebesgue measurable subset of $\mathbb{R}$. Then*

1. $\lambda(A) = \inf\{\lambda(U) : U$ *is open and* $A \subseteq U\}$, *and*
2. $\lambda(A) = \sup\{\lambda(K) : K$ *is compact and* $K \subseteq A\}$.

This theorem is useful in that it tells us that Lebesgue measurable sets can be "approximated from the outside" by open sets and can be "approximated from the inside" by compact sets. This idea is closely related to that of a regular measure, which we will discuss in more detail later. We now look at another useful property of Lebesgue measure concerning its translation invariance. First, if $A \subseteq \mathbb{R}$ and $x \in A$, we define the set $A + x$ as

$$A + x = \{y \in \mathbb{R} : y = a + x \text{ for some } a \in A\}.$$

We call the set $A + x$ the *translate* of A by x.

Theorem 2.3.2. *Let A be a subset of $\mathbb{R}$.*

1. *If $x \in \mathbb{R}$, then $\lambda^*(A) = \lambda^*(A + x)$.*
2. *A is Lebesgue measurable if and only if $A + x$ is Lebesgue measurable.*

One particular set that is a great source of useful examples and counterexamples is the Cantor set. Recall that the Cantor set is defined by constructing a sequence of sets as follows. First define $K_0 = [0,1]$. Now construct K_1 by removing the middle third open interval $\left(\frac{1}{3}, \frac{2}{3}\right)$ from K_0, leaving $K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. We construct K_2 by removing the middle third open interval from each piece of K_1, leaving

$$K_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

We continue this process indefinitely, contructing K_n by removing the middle third open interval from each piece of K_{n-1}, so that K_n is the union of 2^n disjoint closed intervals, each with length $\left(\frac{1}{3}\right)^n$. We define the Cantor set K to be the set of points that remains after this process, so that $K = \cap_{n=0}^{\infty} K_n$. It is a fact from basic real analysis that the Cantor set is compact and has the same cardinality as that of the real numbers [Rud76]. Given these two facts, our next result can be unexpected!

Theorem 2.3.3. *The Cantor set, which is compact and has the cardinality of the continuum, has Lebesgue measure* 0.

Proof. Since each set K_n is the finite union of closed intervals, each K_n must be a Borel set, and hence Lebesgue measurable. Thus, the Cantor set K must also be Lebesgue measurable as the countable intersection of Lebesgue measurable sets, by properties of σ-algebras. We have already noted that Lebesgue measure assigns to each interval its length. First note that $K \subsetneq K_n$ for all $n \geq 0$, and that $\lambda(K_n) = 2^n \left(\frac{1}{3}\right)^n = \left(\frac{2}{3}\right)^n$, since it consists of 2^n disjoint intervals, each of length $\left(\frac{1}{3}\right)^n$.

Since λ is monotone (as a measure), we must have that $\lambda(K) \leq \lambda(K_n) = \left(\frac{2}{3}\right)^n$ for all $n \geq 0$. But λ is also non-negative and $\lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0$, so $\lambda(K) = 0$. □

We may at this point be asking ourselves, since we proved that every Borel set is Lebesgue measurable, is it also true that every subset of $\mathbb{R}$ is Lebesgue measurable? This question is answered with our next theorem.

Theorem 2.3.4. *There exists a subset of* $\mathbb{R}$ *that is not Lebesgue measurable.*

Proof. We first define a relation on $\mathbb{R}$ by letting $x \sim y$ if and only if $x - y$ is a rational number. For any $x \in \mathbb{R}$, $x - x = 0 \in \mathbb{Q}$, so $x \sim x$, and so $\sim$ is reflexive. If $x, y \in \mathbb{Q}$ with $x \sim y$, then $x - y = q \in \mathbb{Q}$ and $y - x = -q$, which is also in $\mathbb{Q}$. Thus $y \sim x$ and $\sim$ is symmetric. Let $x, y, z \in \mathbb{Q}$ with $x \sim y$ and $y \sim z$. So $x - y = r \in \mathbb{Q}$ and $y - z = s \in \mathbb{Q}$. This implies that $x - z = (x - y) + (y - z) = r + s$, which is in $\mathbb{Q}$ since both r and s are. Thus $x \sim z$ and so $\sim$ is transitive. Therefore, $\sim$ is an equivalence relation.

Note that each equivalence class under $\sim$ has the form $\mathbb{Q} + x$ for some $x \in \mathbb{R}$, so that each equivalence class under $\sim$ is dense in $\mathbb{R}$. Since these equivalence classes are disjoint by definition, and since each one must intersect the interval $(0, 1)$, we can now use the axiom of choice to construct a set E that consists of exactly one point from each equivalence class under $\sim$ such that $E \subseteq (0, 1)$. It is this set E that we will prove is not Lebesgue measurable.

Let $\{r_n\}_{n=1}^{\infty}$ be an enumeration of the rational numbers in the interval $(-1, 1)$. For each $n \geq 1$, define $E_n = E + r_n$. We first note that if $e \in E_m \cap E_n$, then there exist e' and e'' in E so that $e = e' + r_m = e'' + r_n$. We can rearrange this equation to obtain $e' - e'' = r_n - r_m \in \mathbb{Q}$, so $e' \sim e''$. But since E has only one element from each equivalence class, and $e' \sim e''$, we must have that $e' = e''$, so $n = m$. Thus, if $m \neq n$, $E_m \cap E_n = \emptyset$, so the sets E_n are disjoint. Also, since $E \subseteq (0, 1)$ and $\{r_n\}_{n=1}^{\infty} \subseteq (-1, 1)$, each set E_n must be included in the interval $(-1, 2)$, so we must have that $\bigcup_{n=1}^{\infty} E_n \subseteq (-1, 2)$. Therefore,

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \lambda\big((-1, 2)\big) = 3.$$

Now, let x be an arbitrary element of $(0, 1)$, and let e be the (unique) element of E so that $x \sim e$. So $x - e \in \mathbb{Q}$ and $x - e$ must also be in $(-1, 1)$, since both x and e are in $(0, 1)$. But $\{r_n\}_{n=1}^{\infty}$ is an enumeration of the rationals in $(-1, 1)$, so $x - e$ must be equal to r_{n_0} for some $n_0 \geq 1$. Thus, $x - e = r_{n_0}$, meaning $x = e + r_{n_0}$, so by definition, $x \in E_{n_0}$. Therefore, $(0, 1) \subseteq \bigcup_{n=1}^{\infty} E_n$, so

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \geq \lambda\big((0, 1)\big) = 1.$$

Now assume that the set E is Lebesgue measurable. Thus, by Theorem 2.3.2, the set E_n is measurable for all $n \geq 1$. Since we proved that the sets E_n are disjoint, this means that

$$\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \lambda(E_n),$$

by the countable additivity of λ. Furthermore, we have that $\lambda(E) = \lambda(E_n)$ for all $n \geq 1$, since $E_n = E + r_n$ and λ is translation invariant, again by Theorem 2.3.2. We now have two cases:

First, if $\lambda(E) = 0$, then

$$\begin{aligned}\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) &= \sum_{n=1}^{\infty} \lambda(E_n) \\ &= \sum_{n=1}^{\infty} \lambda(E) \\ &= \sum_{n=1}^{\infty} 0 \\ &= 0,\end{aligned}$$

a contradiction, since we showed that $\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \geq 1$.

Second, if $\lambda(E) = \alpha \neq 0$, then

$$\begin{aligned}\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) &= \sum_{n=1}^{\infty} \lambda(E_n) \\ &= \sum_{n=1}^{\infty} \lambda(E) \\ &= \sum_{n=1}^{\infty} \alpha \\ &= \infty,\end{aligned}$$

a contradiction, since we showed that $\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) \leq 3$.

So, the assumption that E is Lebesgue measurable leads to a contradiction in all possible cases. Therefore, the set E is not Lebesgue measurable. $\square$

Since, historically, the axiom of choice has sometimes been a subject of disagreement in the field, it was a question among mathematicians whether or not its use was necessary in proving the existence of a subset of $\mathbb{R}$ that was not Lebesgue measurable. Interestingly, this remained an open question until 1970, when Robert M. Solovay published a proof that the existence of such a set cannot be proven from the axioms of Zermelo-Frankel set theory without using the axiom of choice, if a certain consistency assumption holds [Coh97].

We will conclude our discussion of measures with a couple of definitions.

Definition 2.3.5. Let $(X, \mathscr{A}, \mu)$ be a measure space. The measure μ (and also the measure space) is called *complete* if any subset of a measurable set of measure 0 is itself measurable, that is, if whenever we have that $A \in \mathscr{A}$ and $\mu(A) = 0$, then $B \subseteq A$ implies $B \in \mathscr{A}$.

It is a fact that $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda)$ is not a complete measure space, but that the measure space $(\mathbb{R}, \mathscr{A}_{\lambda^*}, \lambda)$ is complete [Coh97]. Our next definition was briefly mentioned earlier in our discussion of the properties of Lebesgue measure.

Definition 2.3.6. Let $\mathscr{A}$ be a σ-algebra on $\mathbb{R}$ that includes the Borel σ-algebra $\mathscr{B}(\mathbb{R})$. A measure μ on $(\mathbb{R}, \mathscr{A})$ is *regular* if

1. for every compact subset K of $\mathbb{R}$, $\mu(K) < \infty$,
2. for every set $A \in \mathscr{A}$,
$$\mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\}, \text{ and}$$
3. for every open set $U \subseteq \mathbb{R}$,
$$\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ and } K \text{ is compact}\}.$$

Condition 2 is sometimes called *outer regularity*, and condition 3 is sometimes called *inner regularity*.

We will now return to something that was mentioned earlier, which is the similarity of the definition of a regular measure to Theorem 2.3.1. In fact, Lebesgue measure is actually a regular measure both on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ and on $(\mathbb{R}, \mathscr{A}_{\lambda^*})$. We will prove this fact in our next theorem.

Theorem 2.3.7. *Lebesgue measure λ is regular on both $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ and $(\mathbb{R}, \mathscr{A}_{\lambda^*})$.*

Proof. If $K \subseteq \mathbb{R}$ is a compact set, then K must be closed and bounded by the Heine-Borel Theorem [Rud76]. Since K is closed, it must be measurable with respect to both $\mathscr{B}(\mathbb{R})$ and $\mathscr{A}_{\lambda^*}$. Since it is bounded, there exists an interval $(a, b) \subseteq \mathbb{R}$, with $a, b \in \mathbb{R}$, so that $K \subseteq (a, b)$. Thus, $\lambda(K) \leq \lambda((a, b)) = b - a < \infty$. Thus, condition 1 of the definition of a regular measure is satisfied.

The fact that λ satisfies conditions 2 and 3 of the definition follow directly from Theorem 2.3.1. Therefore, Lebesgue measure λ is regular with respect to both $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ and $(\mathbb{R}, \mathscr{A}_{\lambda^*})$. $\square$

Chapter 3

Integration

We have now completed our discussion of measures and their properties. Perhaps the most important application of measures is their use in defining the integral in a generalized way. Before we can get to that point, we first need a few more preliminary concepts. We start this chapter by studying measurable functions, and use these concepts to construct the generalized Lebesgue integral with respect to a measure μ. After we present some of the basic properties of integration, we will prove a few extremely useful theorems regarding integration and limits that are part of the foundation for the power of the integral.

3.1 Measurable Functions

Definition 3.1.1. Let $(X, \mathscr{A})$ be a measurable space, and let $A \subseteq X$ with $A \in \mathscr{A}$. A function $f : A \to [-\infty, \infty]$ is called *measurable* with respect to $\mathscr{A}$ if, for every real number t, the set $\{x \in A : f(x) \leq t\}$ is an element of the σ-algebra $\mathscr{A}$.

Note that another way of stating this definition is that f is measurable if, for every real number t, the set $f^{-1}\big((-\infty, t]\big) \in \mathscr{A}$. Consider a set of the form $\{x \in A : f(x) < t\}$. We can express this set in terms of sets of the form in Definition 3.1.1 as

$$\{x \in A : f(x) < t\} = \bigcup_{n=1}^{\infty} \left\{x \in A : f(x) \leq t - \frac{1}{n}\right\}.$$

If f is a measurable function with respect to a measurable space $(X, \mathscr{A})$, this expression implies that, for any real number t, any set of the form $\{x \in A : f(x) < t\}$ is also in $\mathscr{A}$, as a countable union of elements in $\mathscr{A}$.

Now consider a set of the form $\{x \in A : f(x) \geq t\}$. Similarly, we can express this set in terms of sets of the form $\{x \in A : f(x) < t\}$ as

$$\{x \in A : f(x) \geq t\} = A \setminus \{x \in A : f(x) < t\}.$$

Thus, if f is a measurable function with respect to a measurable space $(X, \mathscr{A})$, then for any real number t, any set of the form $\{x \in A : f(x) \geq t\}$ is in $\mathscr{A}$ as the set difference of two elements of A.

Using these and other very similar expressions, we get the following useful theorem.

Theorem 3.1.2. *Let $(X, \mathscr{A})$ be a measurable space, and let $A \in \mathscr{A}$. Let $f : A \to [-\infty, \infty]$ be a function defined on A. Then the following conditions are equivalent:*

1. *f is measurable with respect to $\mathscr{A}$*
2. *for every real number t, the set $\{x \in A : f(x) \leq t\}$ is an element of $\mathscr{A}$;*
3. *for every real number t, the set $\{x \in A : f(x) < t\}$ is an element of $\mathscr{A}$;*
4. *for every real number t, the set $\{x \in A : f(x) \geq t\}$ is an element of $\mathscr{A}$;*
5. *for every real number t, the set $\{x \in A : f(x) > t\}$ is an element of $\mathscr{A}$.*

When the measurable space in question is $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, then a measurable function is called *Borel measurable* or a *Borel function*. When a function is measurable with respect to $(\mathbb{R}, \mathscr{A}_{\lambda^*})$, then the function is called a *Lebesgue measurable* function. Measurable functions play a key role in integration, as we shall see, since the definitions we will use for the integral only apply to measurable functions. Therefore it is of great importance that, before we discuss the integral, we establish some important and useful properties of measurable functions, which will be used later on when we prove things about integration. Many of these properties rely on properties of the underlying σ-algebra, since a function being measurable really has to do with what types of *sets* are guaranteed to be measurable (elements of the σ-algebra). There are two special types of functions that will be fundamental in our construction of the integral, which we will investigate next.

Definition 3.1.3. Let $(X, \mathscr{A})$ be a measurable space, and let $B \subseteq X$. Then the *characteristic function of* B is the function $\chi_B : X \to \mathbb{R}$ defined by

$$\chi_B(x) = \begin{cases} 1, & x \in B \\ 0, & x \notin B \end{cases}.$$

It is a simple consequence of Definition 3.1.1 that χ_B is $\mathscr{A}$-measurable if and only if $B \in \mathscr{A}$.

Definition 3.1.4. A function f is called *simple* if it takes only finitely many values.

Another consequence of Definition 3.1.1 is that if $(X, \mathscr{A})$ is a measurable space, and $f : X \to [-\infty, \infty]$ is a simple function with values $a_1, a_2, \ldots, a_n$, then f is $\mathscr{A}$-measurable if and only $\{x \in X : f(x) = a_i\} \in \mathscr{A}$ for every $i = 1, 2, \ldots, n$. The next theorem lists some useful facts concerning the relationship between two measurable functions with the same domain. The proof of each fact follows from expressions similar to those discussed after the statement of Definition 3.1.1.

Theorem 3.1.5. *Let $(X, \mathscr{A})$ be a measurable space, let $A \in \mathscr{A}$, and let f and g be $[-\infty, \infty]$-valued measurable functions defined on A. Then the following statements are all true:*

1. *the set $\{x \in A : f(x) < g(x)\} \in \mathscr{A}$;*
2. *the set $\{x \in A : f(x) \leq g(x)\} \in \mathscr{A}$;*
3. *the set $\{x \in A : f(x) = g(x)\} \in \mathscr{A}$;*
4. *the maximum of f and g, defined by $(f \vee g)(x) = \max(f(x), g(x))$, is a measurable function from A to $[-\infty, \infty]$;*
5. *the minimum of f and g, defined by $(f \wedge g)(x) = \min(f(x), g(x))$, is a measurable function from A to $[-\infty, \infty]$.*

This theorem is very useful when dealing with two measurable functions, but we will often be working with sequences of measurable functions. We will now prove some facts regarding these types of sequences that will frequently be useful as we continue our investigation of integration.

Theorem 3.1.6. *Let $(X, \mathscr{A})$ be a measurable space, let $A \in \mathscr{A}$, and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of $[-\infty, \infty]$-valued measurable functions defined on A. Then*

1. *the functions $\sup_n f_n$ and $\inf_n f_n$ are measurable functions,*

2. *the functions $\limsup_n f_n$ and $\liminf_n f_n$ are measurable functions, and*

3. *the function $\lim_n f_n$ is a measurable function, where the domain of $\lim_n f_n$ is the set $A_0 = \{x \in A : \limsup_n f_n(x) = \liminf_n f_n(x)\}$ (the set of points at which the limit exists).*

Proof. To prove that $\sup_n f_n$ is measurable, we need to prove that, for any real number t, $\{x \in A : \sup_n f_n(x) \leq t\} \in \mathscr{A}$. Note that

$$\{x \in A : \sup_n f_n(x) \leq t\} = \bigcap_n \{x \in A : f_n(x) \leq t\},$$

which is measurable since each function f_n is, and the countable intersection of measurable sets is measurable. Thus, $\sup_n f_n$ is measurable, by definition.

Similarly, the expression

$$\{x \in A : \inf_n f_n(x) < t\} = \bigcup_n \{x \in A : f_n(x) < t\}$$

implies that the function $\inf_n f_n$ is a measurable function. Therefore we have proven part 1 of the theorem.

Now we define functions g_k and h_k, for every positive integer k, by $g_k(x) = \sup_{n \geq k} f_n(x)$ and $h_k(x) = \inf_{n \geq k} f_n(x)$. Since we have already proven part 1 of the theorem, we can apply it to the functions g_k and h_k and get that each function is measurable, since each is the supremum or the infimum of a sequence of measurable functions.

Now, note that

$$\limsup_n f_n = \inf_k \left(\sup_{n \geq k} f_n \right) = \inf_k g_k,$$

and that

$$\liminf_n f_n = \sup_k \left(\inf_{n \geq k} f_n \right) = \sup_k h_k.$$

Thus, we can again apply part 1 of the theorem to prove that $\limsup_n f_n$ and $\liminf_n f_n$ are measurable, since they are each the respective infimum and supremum of a sequence of measurable functions. Therefore, part 2 of the theorem is proven.

Next, consider that $A_0 = \{x \in A : \limsup_n f_n(x) = \liminf_n f_n(x)\}$. Since we just proved that each of these two functions are measurable, we can apply part 3 of Theorem 3.1.5 to get that $A_0 \in \mathscr{A}$. To prove that $\lim_n f_n$ is measurable, we need to prove that $\{x \in A_0 : \lim_n f_n(x) \leq t\} \in \mathscr{A}$. Consider the identity

$$\{x \in A_0 : \lim_n f_n(x) \leq t\} = A_0 \cap \{x \in A : \limsup_n f_n(x) \leq t\}.$$

Since $A_0 \in \mathscr{A}$ and we already proved that $\limsup_n f_n$ is measurable, we have that the set $A_0 \cap \{x \in A : \limsup_n f_n(x) \leq t\}$, and hence $\{x \in A_0 : \lim_n f_n(x) \leq t\}$, is in $\mathscr{A}$. Therefore, $\lim_n f_n$ is measurable, and the theorem is proven. □

We will now present some results that deal with arithmetic operations on measurable functions.

Theorem 3.1.7. *Let $(X, \mathscr{A})$ be a measurable space, let $A \in \mathscr{A}$, let f and g be real-valued measurable functions on A, and let α be a real number. Then the functions αf, $f + g$, $f - g$, and fg are measurable. Also, the function $\frac{f}{g}$, with domain $\{x \in A : g(x) \neq 0\}$, is measurable.*

If h and k are $[0, \infty]$-valued measurable functions, and β is a non-negative real number, then βh and $h + k$ are measurable.

We will make frequent use of the following functions in our study of integration. If A is a set and f is a $[-\infty, \infty]$-valued function on A, then we define the *positive part* f^+ of f to be the function defined by

$$f^+(x) = \max(f(x), 0),$$

and the *negative part* f^- of f to be the function defined by

$$f^-(x) = -\min(f(x), 0).$$

It follows from the previous theorems that a $[-\infty, \infty]$-valued function f on a measurable space $(X, \mathscr{A})$ is measurable if and only if f^+ and f^- are measurable. Since the function $|f|$ is equal to $f^+ + f^-$, it is a direct consequence of Theorem 3.1.7 that the absolute value of a measurable function is also measurable.

We now will present another very important result that is fundamental to our forthcoming definition of the integral. It lets us represent certain functions in terms of simple functions, which is very desirable since they are so easy to work with.

Theorem 3.1.8. *Let $(X, \mathscr{A})$ be a measurable space, let $A \in \mathscr{A}$, and let f be a $[0,\infty]$-valued measurable function defined on A. Then there exists a sequence $\{f_n\}_{n=1}^{\infty}$ of $[0,\infty)$-valued simple measurable functions defined on A such that, for every $x \in A$,*

1. *the sequence $\{f_n\}$ is non-decreasing, so that $f_1(x) \leq f_2(x) \leq \cdots$, and*
2. *$f(x) = \lim_n f_n(x)$.*

One last theorem regarding measurable functions presents an alternate way of interpreting measurable functions in terms of open, closed, and Borel sets.

Theorem 3.1.9. *Let $(X, \mathscr{A})$ be a measurable space, and let $A \in \mathscr{A}$. If $f : A \to \mathbb{R}$ is a function defined on A, then the following statements are equivalent:*

1. *f is measurable with respect to $\mathscr{A}$;*
2. *for every open set $U \subseteq \mathbb{R}$, $f^{-1}(U) \in \mathscr{A}$;*
3. *for every closed set $C \subseteq \mathbb{R}$, $f^{-1}(C) \in \mathscr{A}$;*
4. *for every Borel subset $B \subseteq \mathbb{R}$, $f^{-1}(B) \in \mathscr{A}$.*

We close this section by introducing one final piece of terminology that will be used extensively throughout the rest of this chapter.

Definition 3.1.10. Let $(X, \mathscr{A}, \mu)$ be a measure space. We say a property P *holds at almost every x in X* if there exists a set $A \in \mathscr{A}$ such that $\{x \in X : P \text{ fails at } x\} \subseteq A$ and $\mu(A) = 0$.

Other ways to say that P holds at almost every x in X is to say that P holds *μ-almost everywhere*, or just *almost everywhere* if the measure μ is clear. Of course, if $(X, \mathscr{A}, \mu)$ is a complete measure space, then by definition, P holds almost everywhere if and only if $\mu(\{x \in X : P \text{ fails at } x\}) = 0$.

3.2 Integration

In this section, we will construct the integral and investigate some of its properties. Our construction will proceed in three steps, each one building off of the last. We first will define the integral for a very small class of functions, and after studying some properties of this integral we will be able to define integration for a larger class of functions. Our goal is to define integration for arbitrary $[-\infty, \infty]$-valued measurable functions, which will be the last step of our construction.

First, we must start with the following. If $(X, \mathscr{A})$ is a measurable space, then we denote the collection of all real-valued simple $\mathscr{A}$-measurable functions on X by $\mathscr{S}$, and the collection of all non-negative functions in $\mathscr{S}$ by $\mathscr{S}_+$. We will start our construction by defining integration over $\mathscr{S}_+$.

Definition 3.2.1. Let μ be a measure on $(X, \mathscr{A})$. Let f belong to $\mathscr{S}_+$ and be given by $f = \sum_{i=1}^{m} a_i \chi_{A_i}$, where $a_1, a_2, \ldots, a_m$ are non-negative real numbers and $A_1, A_2, \ldots, A_m$ are disjoint elements of $\mathscr{A}$ such that f takes the value a_i on the set A_i for every i. Then the *integral of f with respect to μ* is defined to be the sum $\sum_{i=1}^{m} a_i \mu(A_i)$. That is,

$$\int f \, d\mu = \sum_{i=1}^{m} a_i \mu(A_i).$$

Note that this sum is either a real number or ∞. We need to verify that the integral is well-defined, since the representation of the function f is not unique. This is the purpose of our next theorem.

Theorem 3.2.2. *If μ is a measure on the measurable space $(X, \mathscr{A})$, and $f \in \mathscr{S}_+$, then the integral $\int f \, d\mu$ is independent of the representation of f.*

Proof. Let f be given by both $f = \sum_{i=1}^{m} a_i \chi_{A_i}$ and $f = \sum_{j=1}^{n} b_j \chi_{B_j}$, where $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ are non-negative real numbers, $A_1, \ldots, A_m$ are disjoint elements of $\mathscr{A}$, and $B_1, \ldots, B_n$ are disjoint elements of $\mathscr{A}$. Without loss of generality, we can assume that $\bigcup_{i=1}^{m} A_i = \bigcup_{j=1}^{n} B_i$ by disregarding the sets A_i for which $a_i = 0$ and the sets B_j for which $b_j = 0$ if necessary, since they contribute nothing to the value of f at any point $x \in X$.

Now, if $A_i \bigcap B_j \neq \emptyset$, then there exists $x \in A_i \bigcap B_j$ so that $f(x) = a_i$ and $f(x) = b_j$, by the conditions on the sets A_i and B_j. Thus, $a_i = b_j$, since f is well-defined. If $A_i \bigcap B_j = \emptyset$, then $\mu\left(A_i \bigcap B_j\right) = 0$. Also, if i_0 is fixed, then $A_{i_0} = \bigcup_{j=1}^{n} \left(A_{i_0} \cap B_j\right)$, so the countable additivity of the measure μ gives us that

$$a_{i_0} \mu(A_{i_0}) = \sum_{j=1}^{n} a_{i_0} \mu\left(A_{i_0} \cap B_j\right) = \sum_{j=1}^{n} b_j \mu\left(A_{i_0} \cap B_j\right).$$

Using these equalities, we can now apply the definition of the integral of f to get that

$$\begin{aligned}
\int f\,d\mu &= \int\left(\sum_{i=1}^{m} a_i\chi_{A_i}\right)d\mu \\
&= \sum_{i=1}^{m} a_i\mu\left(A_i\right) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} a_i\mu\left(A_i\cap B_j\right) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} b_j\mu\left(A_i\cap B_j\right) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{m} b_j\mu\left(A_i\cap B_j\right) \\
&= \sum_{j=1}^{n} b_j\mu\left(B_j\right) \\
&= \int\left(\sum_{j=1}^{n} b_j\chi_{B_j}\right)d\mu.
\end{aligned}$$

Thus, $\int f\,d\mu$ is the same regardless of which representation is used, so the integral is well-defined and the theorem is proven. □

Our next theorem verifies that some basic properties of the Riemann integral are still valid with our new definition of integration. Particularly, we see that the integral defined above is linear and monotonic. The proof of the first two facts follows from arguments similar to those in the proof of Theorem 3.2.2, and the third follows from the fact that if f and g in $\mathscr{S}_+$ are such that $f(x)\le g(x)$ for all $x\in X$, then the function $g-f$ is also in $\mathscr{S}_+$, and

$$\int g\,d\mu=\int(f+(g-f))\,d\mu=\int f\,d\mu+\int(g-f)\,d\mu\ge\int f\,d\mu.$$

Theorem 3.2.3. *Let $(X,\mathscr{A})$ be a measure space, let f and g be elements of $\mathscr{S}_+$, and let α be a non-negative real number. Then the following statements are true:*

1. $\int\alpha f\,d\mu=\alpha\int f\,d\mu$;
2. $\int(f+g)\,d\mu=\int f\,d\mu+\int g\,d\mu$;
3. *if* $f(x)\le g(x)$ *for all* $x\in X$, *then* $\int f\,d\mu\le\int g\,d\mu$.

Recall that Theorem 3.1.8 stated that for every $[0,\infty]$-valued measurable function f defined on a measurable subset A of a measurable space $(X,\mathscr{A})$, there exists a non-decreasing sequence $\{f_n\}$ of non-negative real-valued simple measurable functions (i.e. elements of $\mathscr{S}_+$) on A so that $f(x) = \lim_n f_n(x)$ for all $x \in X$. We can apply this theorem to elements of $\mathscr{S}_+$ since any function in $\mathscr{S}_+$ is also $[0,\infty]$-valued and measurable, which leads us to the following theorem. A much more general version of this theorem will be proven later on.

Theorem 3.2.4. *Let $(X,\mathscr{A},\mu)$ be a measure space, let $f \in \mathscr{S}_+$, and let $\{f_n\}_{n=1}^{\infty}$ be a non-decreasing sequence of functions in $\mathscr{S}_+$ such that $f(x) = \lim_n f_n(x)$ for all $x \in X$ (note that the existence of such a sequence is guaranteed by Theorem 3.1.8). Then*

$$\int f\,d\mu = \int \lim_n f_n\,d\mu = \lim_n \int f_n\,d\mu.$$

This theorem is very useful in that it gives us conditions for when limits of everywhere-convergent sequences of certain functions and our new integration commute, but it is quite limited (for now) in its scope. So far, we have defined a generalized integral that is based on the ideas of measures and measurable functions. However, since at this point we can only integrate non-negative real-valued simple measurable functions, which is a very limited collection, we are not satisfied and would like to proceed to define the integral for a broader class of functions. Theorem 3.1.8 and Definition 3.2.1 give us some direction, and we can now define integration for arbitrary $[0,\infty]$-valued measurable functions as follows.

Definition 3.2.5. Let $(X,\mathscr{A},\mu)$ be a measure space, and let f be a $[0,\infty]$-valued $\mathscr{A}$-measurable function defined on X. Then the *integral of f with respect to μ* is defined to be

$$\int f\,d\mu = \sup\left\{\int g\,d\mu : g \in \mathscr{S}_+ \text{ and } g \leq f\right\}.$$

If, in the above definition, our function f happens to be an element of $\mathscr{S}_+$ to begin with, then $\sup\left\{\int g\,d\mu : g \in \mathscr{S}_+ \text{ and } g \leq f\right\}$ is just equal to $\int f\,d\mu$ (with respect to Definition 3.2.1), so our new definition of integration agrees with Definition 3.2.1. We will now briefly look at a few properties of this new integral, however, the properties themselves are merely extensions of the properties we have already discussed in relation to integration of elements of $\mathscr{S}_+$. We first present a stronger version of Theorem 3.2.4, which will be strengthened even further in the next section.

Theorem 3.2.6. *Let $(X, \mathscr{A}, \mu)$ be a measure space, let f be a $[0, \infty]$-valued $\mathscr{A}$-measurable function on X, and let $\{f_n\}_{n=1}^{\infty}$ be a non-decreasing sequence of functions in $\mathscr{S}_+$ such that $f(x) = \lim_n f_n(x)$ for all $x \in X$. Then*

$$\int f \, d\mu = \int \lim_n f_n \, d\mu = \lim_n \int f_n d\mu.$$

Theorem 3.2.6 can be used to prove the next theorem regarding some expected properties of our new integral.

Theorem 3.2.7. *Let $(X, \mathscr{A}, \mu)$ be a measure space, let f and g be $[0, \infty]$-valued $\mathscr{A}$-measurable functions on X, and let α be a non-negative real number. Then the following statements are true:*

1. $\int \alpha f \, d\mu = \alpha \int f \, d\mu$;
2. $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$;
3. *if* $f(x) \leq g(x)$ *for all* $x \in X$*, then* $\int f \, d\mu \leq \int g \, d\mu$.

While we now can integrate a much larger class of functions, namely the class of $[0, \infty]$-valued measurable functions, we are now ready to complete the last step of our construction and define integration for arbitrary $[-\infty, \infty]$-valued measurable functions. To do so we must recall the notions of the positive part f^+ and the negative part f^- of an extended real-valued function f. If we have an arbitrary $[-\infty, \infty]$-valued measurable function f, then both f^+ and f^- are $[0, \infty]$-valued measurable functions, and so we can apply Definition 3.2.5 to integrate each part. This leads us to the following definition:

Definition 3.2.8. Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be a $[-\infty, \infty]$-valued $\mathscr{A}$-measurable function defined on X. If at least one of $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are finite, then we say that the integral of f *exists*, and we define the *integral of f with respect to μ* to be

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

Note that this integral may be either ∞ or $-\infty$. However, if *both* $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are finite, then we call the function f *integrable*, and its integral will be a real number. If $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are both infinite, then we say that the integral does not exist. We denote the collection of all real-valued integrable functions on X by

$\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$, which we sometimes abbreviate to just $\mathscr{L}^1$ if the σ-algebra and measure are clear. This definition of integration is commonly called *Lebesgue integration*, whether the measure in question is the Lebesgue measure λ or not, and will be a fundamental part of the rest of this work.

We will now examine some properties of the Lebesgue integral, and we begin by presenting an extension of Theorem 3.2.3. The simple proof of this theorem is found by decomposing the functions f and g into their parts f^+, f^-, g^+, and g^-, and by using the inequalities $(f+g)^+ \leq f^+ + g^+$ and $(f+g)^- \leq f^- + g^-$.

Theorem 3.2.9. *Let $(X, \mathscr{A}, \mu)$ be a measure space, let f and g be real-valued integrable functions on X, and let α be a real number. Then:*

1. *αf and $f+g$ are integrable;*
2. *$\int \alpha f \, d\mu = \alpha \int f \, d\mu$;*
3. *$\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu$;*
4. *if $f(x) \leq g(x)$ for all $x \in X$, then $\int f \, d\mu \leq \int g \, d\mu$.*

The following theorem will be useful when we prove the Dominated Convergence Theorem in the next section; it describes how integration deals with absolute values.

Theorem 3.2.10. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f be a $[-\infty, \infty]$-valued $\mathscr{A}$-measurable function on X. Then f is integrable if and only if $|f|$ is integrable. Furthermore, if f and $|f|$ are integrable, then*

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

Proof. By definition, $|f|$ is integrable if and only if $\int |f|^+ \, d\mu$ and $\int |f|^- \, d\mu$ are both finite. Since $|f| = f^+ + f^-$, and each of these are non-negative functions, then $|f|^+ = f^+ + f^-$ and $|f|^- = 0$. Since the 0 function is simple and can be expressed as $0\chi_{\emptyset}$, we can use our first definition of integration to get that $\int 0 \, d\mu = 0\mu(\emptyset) = 0 < \infty$, so $|f|^- = 0$ is always finite with respect to any measure. Therefore, $|f|$ is integrable if and only if $\int (f^+ + f^-) \, d\mu$ is finite.

If f is integrable, then $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are both finite. Thus, by Theorem 3.2.7, $\int (f^+ + f^-) \, d\mu = \int f^+ \, d\mu + \int f^- \, d\mu$ is also finite, so $|f|$ is integrable. If we assume that $|f|$ is integrable, then $\int(f^+ + f^-)\,d\mu = \int f^+\,d\mu + \int f^-\,d\mu$ is finite, and so, since each is a positive number, we must have that each of $\int f^+\,d\mu$ and $\int f^-\,d\mu$ are finite, so f is integrable. Therefore, f is integrable if and only if $|f|$ is integrable.

Now, if f and $|f|$ are integrable, then we can use the triangle inequality and Theorem 3.2.7 to get that

$$\begin{aligned}
\left|\int f\,d\mu\right| &= \left|\int f^+\,d\mu - \int f^-\,d\mu\right| \\
&\leq \left|\int f^+\,d\mu\right| + \left|\int f^-\,d\mu\right| \\
&= \int f^+\,d\mu + \int f^-\,d\mu \\
&= \int (f^+ + f^-)\,d\mu \\
&= \int |f|\,d\mu.
\end{aligned}$$

Thus, $|\int f\,d\mu| \leq \int |f|\,d\mu$. □

The following theorem illustrates how our new integration essentially "ignores" μ-negligible sets (sets that are contained in sets of measure 0). It will be useful in the proofs of the limit theorems in the next section, and lets us work with functions that meet a certain requirement almost everywhere instead of everywhere, allowing us a larger degree of generalization.

Theorem 3.2.11. *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let f and g be $[-\infty,\infty]$-valued $\mathscr{A}$-measurable functions on X that agree almost everywhere. If either $\int f\,d\mu$ or $\int g\,d\mu$ exists, then both must exist, and $\int f\,d\mu = \int g\,d\mu$.*

Proof. We start by first considering the case where f and g are both $[0,\infty]$-valued. Let $A = \{x \in X : f(x) \neq g(x)\}$ be the set of points at which f and g do not agree. Then A is measurable by part 3 of Theorem 3.1.5, since $A = \{x \in X : f(x) = g(x)\}^c$. Since $f = g$ almost everywhere, we must have that $\mu(A) = 0$.

Now construct a sequence $\{h_n\}_{n=1}^{\infty}$ of functions by defining, for all $n \geq 1$, $h_n(x) = n\chi_A$. Then $\{h_n\}$ is a non-decreasing sequence of non-negative simple real-valued measurable functions, and for each n, $\int h_n\,d\mu = n\mu(A) = 0$. Furthermore, if $x \in A$, then $\lim_{n\to\infty} h_n(x) = \infty$, and if $x \notin A$, then $\lim_{n\to\infty} h_n(x) = 0$. So, if we define a function

$h : X \to [0, \infty]$ by

$$h(x) = \begin{cases} \infty & \text{if } x \in A, \\ 0 & \text{if } x \notin A, \end{cases}$$

then $\lim_n h_n = h$. Therefore, we can apply Theorem 3.2.6 to get that

$$\int h \, d\mu = \lim_n \int h_n \, d\mu = \lim_n 0 = 0.$$

So, since $f \leq g + h$, Theorem 3.2.7 tells us that

$$\int f \, d\mu \leq \int (g + h) \, d\mu = \int g \, d\mu + \int h \, d\mu = \int g \, d\mu.$$

Similarly, since $g \leq f + h$, we also have that $\int g \, d\mu \leq \int f \, d\mu$. Thus, $\int f \, d\mu = \int g \, d\mu$.

Now let f and g be $[-\infty, \infty]$-valued functions. Recall that $f = f^+ - f^-$ and $g = g^+ - g^-$. Since $f = g$ almost everywhere, we must have that $f^+ = g^+$ and $f^- = g^-$ almost everywhere as well. Since each of these functions are $[0, \infty]$-valued, we can apply the above result to see that $\int f^+ \, d\mu = \int g^+ \, d\mu$ and $\int f^- \, d\mu = \int g^- \, d\mu$. Therefore,

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu = \int g^+ \, d\mu - \int g^- \, d\mu = \int g \, d\mu. \qquad \square$$

Our last theorem for this section summarizes some interesting facts regarding integration and properties that hold almost everywhere.

Theorem 3.2.12. *Let $(X, \mathscr{A}, \mu)$ be a measure space and let f be a $[-\infty, \infty]$-valued $\mathscr{A}$-measurable function defined on X. Then the following statements are true:*

1. *If $\int |f| \, d\mu = 0$, then $f = 0$ almost everywhere.*

2. *If f is integrable, then $|f| < \infty$ almost everywhere.*

3. *f is integrable if and only if there exists a function f' in $\mathscr{L}^1(X, \mathscr{A}, \mu, \mathbb{R})$ such that $f = f'$ almost everywhere.*

3.3 Integral Limit Theorems

In this section, we will present some powerful theorems regarding limits and the Lebesgue integral. These results do not apply to Riemann integration, which is one of the reasons Lebesgue integration is so useful. The following theorems give conditions under which limits and integration can commute when working with convergent sequences of measurable functions. Once we have proven them, we will have some powerful tools at our disposal that "may well be regarded as the core of the Lebesgue theory" [Rud76]. The first of these is the generalization of Theorem 3.2.4 and Theorem 3.2.6 to the Lebesgue integral.

Theorem 3.3.1 (The Monotone Convergence Theorem). *Let $(X, \mathscr{A}, \mu)$ be a measure space. Let f and $f_1, f_2, \ldots$ be $[0, \infty]$-valued $\mathscr{A}$-measurable functions on X such that*

$$f_1(x) \leq f_2(x) \leq \cdots$$

and

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

both hold at almost every $x \in X$. Then

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

Proof. We will first consider the case where the relations in the statement of the theorem hold for every $x \in X$ rather than almost everywhere. Since the sequence $\{f_n\}_{n=1}^{\infty}$ is non-decreasing, then it must be that $f_n \leq f$ for every $n \geq 1$. Therefore, by Theorem 3.2.7,

$$\int f_1\, d\mu \leq \int f_2\, d\mu \leq \cdots \leq \int f\, d\mu.$$

Thus, the sequence of integrals $\{\int f_n\, d\mu\}_{n=1}^{\infty}$ is monotonically increasing with upper bound $\int f d\mu$. Therefore, the sequence $\{\int f_n\, d\mu\}$ must converge to a number in $[0, \infty]$ (which may or may not be ∞), and the limit of this sequence is such that

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

Now, for every $n \geq 1$, we use Theorem 3.1.8 to choose a non-decreasing sequence $\{g_{n,k}\}_{k=1}^{\infty}$ of non-negative real-valued measurable simple functions defined on X such that $f_n = \lim_{k\to\infty} g_{n,k}$. Also for every $n \geq 1$, define a function $h_n : X \to [0, \infty)$ by $h_n(x) = \max(g_{1,n}(x), g_{2,n}(x), \ldots, g_{n,n}(x))$.

Note that $\{h_n\}_{n=1}^{\infty}$ is clearly a non-decreasing sequence, and each h_n is simple, measurable, and takes values in $[0, \infty)$. Also, by definition of the functions h_n, for every

$n \geq 1$ we have that $h_n \leq f_n$ and $f = \lim_{n\to\infty} h_n$. Therefore, we can apply Theorem 3.2.6 and Theorem 3.2.7 to find that

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

Since we have already proved the reverse inequality, we have that $\int f\, d\mu = \lim_n \int f_n\, d\mu$.

Now assume that the relations in the statement of the theorem only hold at almost every $x \in X$. Let A be a set such that

$$\Big(\{x \in X : f_n(x) > f_{n+1}(x) \text{ for some } n \geq 1\} \cup \{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\}\Big) \subseteq A$$

and so that $\mu(A) = 0$.

Consider the function $f\chi_{A^c}$ and the sequence of functions $\{f_n\chi_{A^c}\}_{n=1}^{\infty}$. Since $f_1\chi_{A^c}(x) \leq f_2\chi_{A^c}(x) \leq \cdots$ for all $x \in A^c$ and $f_n\chi_{A^c}(x) = 0$ for all $n \geq 1$ and $x \in A$, we must have that $f_1\chi_{A^c}(x) \leq f_2\chi_{A^c}(x) \leq \cdots$ at every $x \in X$. Similarly, $f\chi_{A^c}(x) = \lim_n f_n\chi_{A^c}(x)$ at every $x \in X$. Therefore we can apply the results of the first part of this proof to get that

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

Now, since $\mu(A) = 0$, we must have that, for every $n \geq 1$, $f_n\chi_{A^c} = f_n$ almost everywhere and $f\chi_{A^c} = f$ almost everywhere. Thus, we can use Theorem 3.2.11 to get that $\int f_n\chi_{A^c}\, d\mu = \int f_n\, d\mu$ and that $\int f\chi_{A^c}\, d\mu = \int f\, d\mu$. Combining these equalities with the above equation gives us that

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

Our next result will be used in our proof of the Dominated Convergence Theorem, but even though the name by which it is commonly known contains the word "lemma," it is presented here as a theorem due to its wide applicability beyond just that particular proof.

Theorem 3.3.2 (Fatou's Lemma). *Let $(X, \mathscr{A}, \mu)$ be a measure space, and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of $[0, \infty]$-valued $\mathscr{A}$-measurable functions on X. Then*

$$\int \liminf_{n\to\infty} f_n\, d\mu \leq \liminf_{n\to\infty} \int f_n\, d\mu.$$

Proof. Construct a sequence $\{g_n\}_{n=1}^{\infty}$ of functions by defining, for every $n \geq 1$, $g_n = \inf_{k \geq n} f_k$. Therefore, by Theorem 3.1.6, each function g_n is measurable. Since

$$\{f_{n+1}, f_{n+2}, \ldots\} \subseteq \{f_n, f_{n+1}, \ldots\},$$

then by definition of the functions g_n, we must have that $g_1(x) \leq g_2(x) \leq \cdots$ at every $x \in X$. This means that, at every $x \in X$,

$$\lim_{n \to \infty} g_n(x) = \sup_{n \geq 1} (g_n(x)) = \sup_{n \geq 1} \left(\inf_{k \geq n} f_k(x) \right) = \liminf_{n \to \infty} f_n(x).$$

Now, since we have that $g_n \leq f_n$ for all $n \geq 1$, we can apply the Monotone Convergence Theorem (Theorem 3.3.1) to get that

$$\int \liminf_{n \to \infty} f_n \, d\mu = \int \lim_{n \to \infty} g_n \, d\mu = \lim_{n \to \infty} \int g_n \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu. \qquad \square$$

The last theorem we will present in this section is very useful, as it can be used to show that a particular function is integrable or to even provide an upper bound for the value of the integral of a particular integrable function. This result is perhaps the most powerful limit theorem we will consider.

Theorem 3.3.3 (Lebesgue's Dominated Convergence Theorem). *Let $(X, \mathscr{A}, \mu)$ be a measure space, let g be a $[0, \infty]$-valued integrable function on X, and let f and $f_1, f_2, \ldots$ be $[-\infty, \infty]$-valued $\mathscr{A}$-measurable functions on X such that*

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

and

$$|f_n(x)| \leq g(x) \text{ for all } n \geq 1,$$

both holding at almost every $x \in X$. Then f and $f_1, f_2, \ldots$ are integrable, and

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

Proof. First, let A be a set such that

$$\left(\{x \in X : \lim_{n \to \infty} f_n(x) \neq f(x)\} \cup \{x \in X : |f_n(x)| > g(x) \text{ for some } n \geq 1\} \right) \subseteq A$$

and so that $\mu(A) = 0$. Now construct a sequence of functions $\{h_n\}_{n=1}^{\infty}$ by defining $h_n(x) = f_n \chi_{A^c}(x)$ for all $n \geq 1$, and define a function h by setting $h(x) = f \chi_{A^c}(x)$.

Therefore, $|h_n(x)| = 0 \leq g(x)$ for all $x \in A$, and $|h_n(x)| = |f_n(x)| \leq g(x)$ for all $x \in A^c$, so $|h_n| \leq g$ everywhere, for all $n \geq 1$. Also, by our definition of h, $h = \lim_{n\to\infty} h_n$ everywhere, so that $|h| \leq g$ everywhere as well.

Thus, since $|h_n|$ and $|h|$ are non-negative, we can apply Theorem 3.2.7 to get that $\int |h|\, d\mu \leq \int g\, d\mu$ and that $\int |h_n|\, d\mu \leq \int g\, d\mu$ for all $n \geq 1$. But $f = h$ almost everywhere, and for all $n \geq 1$, $f_n = h_n$ almost everywhere, so by Theorem 3.2.11, for all $n \geq 1$,

$$\int |f_n|\, d\mu = \int |h_n|\, d\mu \leq \int g\, d\mu < \infty$$

and

$$\int |f|\, d\mu = \int |h|\, d\mu \leq \int g\, d\mu < \infty.$$

So, by Theorem 3.2.7, we have that $\int |f_n|\, d\mu = \int {f_n}^+\, d\mu + \int {f_n}^-\, d\mu < \infty$, so each integral must be finite, which means each f_n is integrable, by definition. Similarly, $\int |f|\, d\mu = \int f^+\, d\mu + \int f^-\, d\mu < \infty$, so each integral is finite, and thus f is integrable, by definition.

We have proved the integrability of f and of the functions f_n, and we now turn to the desired equality. We start with the case where the conditions of the theorem hold at every $x \in X$, rather than almost everywhere, and where $g(x) < \infty$ at every $x \in X$. Since each function $|f_n|$ is dominated everywhere by g, then $\{g + f_n\}_{n=1}^{\infty}$ is a sequence of non-negative $\mathscr{A}$-measurable functions. Furthermore, for every $x \in X$,

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

Therefore, since the limit exists at every $x \in X$, we have that $\lim_n (g + f_n) = \liminf_n (g+f_n) = g+f$, and so we can apply Theorem 3.2.9 and Fatou's Lemma (Theorem 3.3.2) to get that

$$\begin{aligned}
\int g\, d\mu + \int f\, d\mu &= \int (g + f)\, d\mu \\
&= \int \liminf_{n\to\infty} (g + f_n)\, d\mu \\
&\leq \liminf_{n\to\infty} \int (g + f_n)\, d\mu \\
&= \liminf_{n\to\infty} \left(\int g\, d\mu + \int f_n\, d\mu \right) \\
&= \int g\, d\mu + \liminf_{n\to\infty} \int f_n\, d\mu.
\end{aligned}$$

By subtracting the finite number $\int g\,d\mu$ from both sides of the above inequality, we obtain that

$$\int f\,d\mu \leq \liminf_{n\to\infty} \int f_n\,d\mu.$$

Similarly, $\{g-f_n\}_{n=1}^{\infty}$ is a sequence of non-negative $\mathscr{A}$-measurable functions such that $\lim_n (g-f_n)(x) = (g-f)(x)$ at every $x \in X$, and so $\lim_n (g-f_n) = \liminf_n (g-f_n) = g-f$. Applying Theorem 3.2.9 and Fatou's Lemma (Theorem 3.3.2) as before gives us that

$$\begin{aligned}
\int g\,d\mu - \int f\,d\mu &= \int (g-f)\,d\mu \\
&= \int \liminf_{n\to\infty}(g-f_n)\,d\mu \\
&\leq \liminf_{n\to\infty} \int (g-f_n)\,d\mu \\
&= \liminf_{n\to\infty} \left(\int g\,d\mu - \int f_n\,d\mu \right) \\
&= \int g\,d\mu + \liminf_{n\to\infty} \left(- \int f_n\,d\mu \right) \\
&= \int g\,d\mu - \limsup_{n\to\infty} \int f_n\,d\mu.
\end{aligned}$$

Rearranging the above inequality shows that

$$\limsup_{n\to\infty} \int f_n\,d\mu \leq \int f\,d\mu.$$

Therefore, we have that

$$\limsup_{n\to\infty} \int f_n\,d\mu \leq \int f\,d\mu \leq \liminf_{n\to\infty} \int f_n\,d\mu,$$

but $\liminf_n \int f_n\,d\mu$ is always less than or equal to $\limsup_n \int f_n\,d\mu$, so it must be that

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

Now we consider the case where the conditions of the theorem hold only almost everywhere, and where $g(x) < \infty$ is not necessarily true at every $x \in X$. Note, however, since g is integrable by hypothesis, that $|g| = g < \infty$ almost everywhere, by Theorem 3.2.12. Similar to the beginning of this proof, we will now let A be a set containing all points $x \in X$ at which at least one of the following statements is true:

1. $f(x) \neq \lim_n f_n(x)$;

2. $|f_n(x)| > g(x)$ for some $n \geq 1$;

3. $g(x) = \infty$,

and so that $\mu(A) = 0$.

Thus, if we define the sequence of functions $\{h_n\}_{n=1}^{\infty}$ by $h_n(x) = f_n\chi_{A^c}(x)$, the function h by $h(x) = f\chi_{A^c}(x)$, and the function g' by $g'(x) = g\chi_{A^c}(x)$, then we have the following, which all follow from the properties of the functions f_n, f, and g, and the definition of the set A:

1. g' is a $[0, \infty]$-valued integrable function;

2. h and each h_n are $[-\infty, \infty]$-valued measurable functions;

3. $h(x) = \lim_n h_n(x)$ at every $x \in X$;

4. $|h_n(x)| \leq g'(x)$ at every $x \in X$, for every $n \geq 1$;

5. $g'(x) < \infty$ at every $x \in X$.

Therefore, we can apply the results of the first part of this proof to get that $\int h\, d\mu = \lim_n \int h_n\, d\mu$. But, also by the definition of the set A, we see that $h = f$ almost everywhere and, for every $n \geq 1$, $h_n = f_n$ almost everywhere. Thus, by Theorem 3.2.11, $\int h\, d\mu = \int f\, d\mu$ and, for every $n \geq 1$, $\int h_n\, d\mu = \int f_n\, d\mu$. Combining these equations gives us that

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

Chapter 4

The Riesz Representation Theorem for Linear Functionals

In this chapter, we will state and prove the main theorem of this project. In the first section, we will focus on the prerequisite material needed for the statement of the theorem and the preliminary lemmas that we will be using in the proof. The next section will introduce some questions that motivated the work that led to the Riesz representation theorem, and the theorem itself will be presented. The third and final section of this chapter will be devoted entirely to the lengthy proof, which will be broken up into a few pieces in order to help with organization and clarity of presentation.

4.1 Preliminary Material

We begin our preliminary work by recalling some fundamental definitions from topology. Let X be a topological space. First, we say a subset K of X is *compact* if every open cover $\{U_\alpha\}_{\alpha\in\Lambda}$, where Λ is some indexing set, has a finite subcover $\{U_{\alpha_i}\}_{i=1}^{n}$. The space X is called *locally compact* if every point x in X has an open neighborhood U_x such that the closure $\overline{U_x}$ of U_x is a compact set. Lastly, we call the space X *Hausdorff* if, for every pair of distinct points x and y in X, there exist disjoint open sets U and V such that $x \in U$ and $y \in V$. Measures on locally compact Hausdorff spaces have been the subject of much research and discovery, an example of which is the main theorem of this project. Perhaps the reason for this is that, while the conditions imposed upon them are lenient enough to allow results to be widely applicable, locally compact Hausdorff spaces are one of the "most well-behaved classes of spaces to deal with in mathematics" [Mun06]. The following useful results regarding compactness can be used to establish some basic facts concerning measures defined on locally compact Hausdorff spaces. The first one we will examine tells us that, in a Hausdorff space, the "Hausdorff-ness" property extends to compact sets as well, in the sense that disjoint compact sets can be separated by open sets.

Theorem 4.1.1. *Let X be a Hausdorff space, and let K and L be disjoint compact subsets of X. Then there exist disjoint open subsets U and V of X so that $K \subseteq U$ and $L \subseteq V$.*

Proof. Firstly, consider the case where one (or both) of our compact sets is the empty set. Without loss of generality, assume that $K = \emptyset$. Then we can choose $U = \emptyset$ and $V = X$, and we see that $U \cap V = \emptyset$, $K = \emptyset \subseteq \emptyset$, and $L \subseteq X$, so the theorem holds.

Now consider the case where K contains exactly one point $x \in X$, so that $K = \{x\}$. Since X is Hausdorff, then for every point $y \in L$ we can choose a pair of disjoint open sets U_y and V_y so that $x \in U_y$ and $y \in V_y$. This implies that the collection $\{V_y\}_{y \in L}$ is an open cover of the compact set L, so there must exist a finite subcover. Thus, there exist points $y_1, y_2, \ldots, y_n \in L$ so that $L \subseteq \bigcup_{i=1}^n V_{y_i}$. Now, let $U = \bigcap_{i=1}^n U_{y_i}$ and $V = \bigcup_{i=1}^n V_{y_i}$. So each of U and V must be open as the finite union or intersection of open sets, and $U \cap V = \emptyset$ by our choice of U and V. Also, $L \subseteq V$, and since x is in each U_y, we must have that $\{x\} = K \subseteq U$. Therefore, we have proven the theorem in the case where K consists of exactly one point.

Now let K have more than one element. By above, for every $x \in K$ there exist disjoint open subsets U_x and V_x of X such that $x \in U_x$ and $L \subseteq V_x$. Therefore, the collection $\{U_x\}_{x \in K}$ is an open cover of the compact set K, so there must exist $x_1, x_2, \ldots, x_m \in K$ such that $K \subseteq \bigcup_{j=1}^m U_{x_j}$. So, similar to the previous arguments, let $U = \bigcup_{j=1}^m U_{x_j}$ and $V = \bigcap_{j=1}^m V_{x_j}$. Thus, U and V are open sets, $K \subseteq U$ and, since L is contained in each V_x, we have that $L \subseteq V$. Also, $U \cap V$ is empty by our choice of U and V, so the theorem is proven for this final case. □

We can use the previous result to prove the following theorem about open neighborhoods of points in locally compact Hausdorff spaces and its important corollary. These facts are used to obtain a crucial theorem in the development of the study of measures on locally compact Hausdorff spaces, which will be an essential part of our proof of the Riesz representation theorem later on.

Theorem 4.1.2. *Let X be a locally compact Hausdorff space, let $x \in X$, and let U be an open neighborhood of x, that is, U is open and $x \in U$. Then there there exists an open neighborhood V of x whose closure $\overline{V}$ is compact and such that $\overline{V} \subseteq U$.*

Proof. By definition of X being locally compact, there exists an open neighborhood W of x whose closure $\overline{W}$ is compact. Since $x \in W$ and $x \in U$, $W \cap U$ is also an open neighborhood of x, and since $\overline{W \cap U}$ is a closed subset of the compact set $\overline{W}$, $\overline{W \cap U}$ must also be compact. Thus, without loss of generality, we can assume that $W \subseteq U$. However, our theorem is not yet proven, since there is no guarantee that $\overline{W} \subseteq U$. To get around this problem, consider the set $\overline{W} \setminus W = \overline{W} \cap W^c$. Since W is open, W^c must be closed, so that $\overline{W} \setminus W$ is closed as the intersection of two closed sets, and is compact as a closed subset of the compact set $\overline{W}$. Also, $x \notin \overline{W} \setminus W$ since $x \in W$, and $\{x\}$ is a compact subset of X since it is finite.

Thus, $\{x\}$ and $\overline{W} \setminus W$ are disjoint compact subsets of X, and so we can use Theorem 4.1.1 to choose disjoint open sets V_1 and V_2 so that $\{x\} \subseteq V_1$ and $\overline{W} \setminus W \subseteq V_2$. Now consider the set $V_1 \cap W$. Since V_1 and W are both open and both contain the point x, $V_1 \cap W$ is an open neighborhood of x. Since $\overline{V_1 \cap W}$ is a closed subset of the compact set $\overline{W}$, $\overline{V_1 \cap W}$ is compact. Also note that, by our choice of V_1 and V_2, it must be true that $\overline{V_1} \subseteq W$. Therefore, we see that

$$\overline{V_1 \cap W} \subseteq \overline{V_1} \cap \overline{W} \subseteq \overline{V_1} \subseteq W \subseteq U,$$

so the set $V_1 \cap W$ is the required set, and our theorem is proven. □

Corollary 4.1.3. *Let X be a locally compact Hausdorff space, let K be a compact subset of X, and let U be an open subset of X so that $K \subseteq U$. Then there exists an open subset V of X whose closure $\overline{V}$ is compact and so that $K \subseteq V \subseteq \overline{V} \subseteq U$.*

Proof. For each point x in K, we can use Theorem 4.1.2 to choose an open neighborhood W_x of x so that $\overline{W_x}$ is closed and $\overline{W_x} \subseteq U$. Thus, the collection $\{W_x\}_{x \in K}$ is an open cover of the compact set K, and so there must exist a finite subcover. Let $x_1, x_2, \ldots, x_n \in K$ be such that $\{W_{x_i}\}_{i=1}^n$ is a finite open cover of K, and define the set $V = \bigcup_{i=1}^n W_{x_i}$. So $K \subseteq V$ by definition, and since any set is always contained in its own closure, we have that $V \subseteq \overline{V}$. Note that $\overline{V} = \overline{(\bigcup_{i=1}^{n} W_{x_i})} = \bigcup_{i=1}^{n} \overline{W_{x_i}}$, since the closure of the union of finitely many sets is equal to the union of the closures of the individual sets. Thus, since each set W_x was chosen so that $\overline{W_x}$ is contained in U, we have that each set $\overline{W_{x_i}} \subseteq U$ and so $\overline{V} \subseteq U$. □

We will now turn to some fundamental concepts involving functions on topological spaces that are essential parts of the Riesz representation theorem. Let X be a topological space, and let $f : X \to \mathbb{R}$ be a continuous real-valued function defined on X. We define the *support* of f to be the closure of the set $\{x \in X : f(x) \neq 0\}$, denoted by $\text{supp}(f)$. If, additionally, X is locally compact and Hausdorff, we denote the set of all continuous real-valued functions $f : X \to \mathbb{R}$ for which $\text{supp}(f)$ is compact by $\mathscr{K}(X)$. The following fact about $\mathscr{K}(X)$ will be necessary in our proof of the Riesz representation theorem. The full proof, which is ommitted here, relies on Corollary 4.1.3 and a powerful theorem from topology known as Urysohn's lemma [Coh97]. The Urysohn lemma has been called by some the first non-trivial result of point-set topology in that its proof is extremely non-obvious, and indeed that "it would take considerably more originality than most of us possess to prove this lemma unless we were given copious hints" [Mun06]. Two variations of the proof of Urysohn's lemma, both depending on the same crucial and brilliant idea, can be found in their entirety in [Mun06] and [Rud87]. For now, we present one of its consequences that will be of use to us later on.

Theorem 4.1.4. *Let X be a locally compact Hausdorff space, let K be a compact subset of X, and let U be an open subset of X such that $K \subseteq U$. Then there exists a function $f \in \mathscr{K}(X)$ such that $\chi_K \leq f \leq \chi_U$ and* $\text{supp}(f) \subseteq U$.

The following consequence of Theorem 4.1.4 will also be necessary in our proof of the main theorem of this project. The proof can be obtained by using Theorem 4.1.1, Theorem 4.1.4, and induction on the number n of open sets.

Theorem 4.1.5. *Let X be a locally compact Hausdorff space, let $f \in \mathscr{K}(X)$, and let $U_1, U_2, \ldots, U_n$ be finitely many open subsets of X such that*

$$\text{supp}(f) \subseteq \bigcup_{i=1}^{n} U_i.$$

Then there exist functions $f_1, f_2, \ldots, f_n$ in $\mathscr{K}(X)$ such that $f = f_1 + f_2 + \cdots + f_n$ and *such that, for every $i \in \{1, 2, \ldots, n\}$,* $\operatorname{supp}(f_i) \subseteq U_i$. *Furthermore, if the function f is non-negative, then each function f_i can be chosen to be non-negative as well.*

Now, recall that the Borel σ-algebra on $\mathbb{R}$ was defined to be the σ-algebra on $\mathbb{R}$ generated by the collection of open subsets of $\mathbb{R}$. In a general topological space, the notion of open sets is generalized, and any subset that is an element of the topology is called open. This leads us to the following definition, which generalizes $\mathscr{B}(\mathbb{R})$ to the Borel σ-algebra over a topological space, rather than just the real numbers.

Definition 4.1.6. Let X be a Hausdorff topological space. Then we define the *Borel σ-algebra* on X to be the σ-algebra generated by the open subsets of X, denoted by $\mathscr{B}(X)$.

We call the elements of $\mathscr{B}(X)$ *Borel sets*, and we call any measure defined on $\mathscr{B}(X)$ a *Borel measure*. Note that it is an elementary consequence of the properties of σ-algebras that $\mathscr{B}(X)$ is also generated by the closed subsets of X, since the complement of any open set is closed. In our study of the Riesz representation theorem, we will be working with Borel measures that satisfy a certain property that is an extension of Definition 2.3.6. This property is our next definition.

Definition 4.1.7. Let X be a Hausdorff topological space, and let μ be a Borel measure defined on $\mathscr{B}(X)$. We say that μ is *regular* if

1. for every compact subset K of X, $\mu(K) < \infty$,

2. for every set $A \in \mathscr{B}(X)$,

$$\mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\}, \text{ and}$$

3. for every open set $U \subseteq X$,

$$\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ and } K \text{ is compact}\}.$$

As with Definition 2.3.6, we sometimes refer to condition 2 as *outer regularity* and condition 3 as *inner regularity*. Regular Borel measures can be very useful, and are the subject of the Riesz representation theorem. Recall that a *linear functional* is a linear map defined on a vector space whose values lie in the field over which the vector space is defined. The power of the Riesz representation theorem is in that it describes how certain linear functionals can be represented by regular Borel measures. We are now ready to investigate the main theorem of this project.

4.2 The Riesz Representation Theorem

In this section, we will investigate the ideas and the natural questions stemming from these ideas that motivated the discovery of the Riesz representation theorem, and we will close the section by stating the theorem itself. A major part of this process was the surprising notion that certain linear functionals seem to correspond with regular Borel measures. To start with, let us present the following fact about the set $\mathscr{K}(X)$.

If X is a locally compact Hausdorff space, recall that $\mathscr{K}(X)$ is defined to be the set of all continuous real-valued functions on X with compact support. Consider the following: If f and g are in $\mathscr{K}(X)$, and k is a real number, then $f+g$ and kf are also real-valued, continuous, and have compact support, so are also in $\mathscr{K}(X)$. The zero function is continuous and real-valued, and $\{x \in X : 0(x) \neq 0\} = \emptyset$, which is closed and compact, so the zero function is in $\mathscr{K}(X)$. If f is in $\mathscr{K}(X)$, then $-f$ is also continuous and real valued. Also, $\{x \in X : f(x) \neq 0\} = \{x \in X : -f(x) \neq 0\}$, so f has compact support if and only if $-f$ does. Thus, $f \in \mathscr{K}(X)$ if and only if $-f \in \mathscr{K}(X)$. Finally, the associativity and commutativity of function addition, the associativity of scalar multiplication with field multiplication, the distributivity of scalar multiplication with respect to vector and field addition, and the existence of the scalar multiplicative identity all follow from the fact that the functions in $\mathscr{K}(X)$ are real-valued. These facts show that the set $\mathscr{K}(X)$ is actually a vector space. The next theorem presents an important fact about the members of $\mathscr{K}(X)$ that will be essential to how we use the vector space.

Theorem 4.2.1. *Let X be a locally compact Hausdorff space, let μ be a regular Borel measure defined on $\mathscr{B}(X)$, and let the function $f : X \to \mathbb{R}$ be a member of $\mathscr{K}(X)$. Then f is integrable with respect to μ.*

Proof. We first must verify that f is Borel measurable. Since $f : X \to \mathbb{R}$ is in $\mathscr{K}(X)$, f must be continuous, so if U is an open subset of $\mathbb{R}$, then $f^{-1}(U)$ is open in X. Thus, $f^{-1}(U) \in \mathscr{B}(X)$, and so f is measurable by Theorem 3.1.9.

Now consider that the continuous function f is only non-zero on a compact subset of X, and recall that continuous functions defined on compact sets are bounded. Thus, it follows that the function f must be bounded. Let $A = \overline{\{x \in X : f(x) \neq 0\}}$ (note that A is compact since $f \in \mathscr{K}(X)$), and let M be a real number such that $|f(x)| < M$ for all $x \in X$. Then $|f| \leq M\chi_A$, so by Theorem 3.2.9,

$$\int |f|\, d\mu \leq \int M\chi_A\, d\mu = M\mu(A) < \infty,$$

since μ is regular and so $\mu(A)$ is finite. Therefore, $|f|$ is integrable with respect to μ, and so by Theorem 3.2.10, f is integrable as well. □

So, if X is a locally compact Hausdorff space and μ is a regular Borel measure on $\mathscr{B}(X)$, let us now consider the map defined on $\mathscr{K}(X)$ by $f \mapsto \int f\, d\mu$. Since each $f \in \mathscr{K}(X)$ is integrable with respect to μ by the Theorem 4.2.1, we have that this map sends elements of the vector space $\mathscr{K}(X)$ to elements of the field $\mathbb{R}$ over which $\mathscr{K}(X)$ is defined, and the map is linear since integration is linear by Theorem 3.2.9. Thus, we see that integration with respect to the regular Borel measure μ is actually a linear functional acting on the vector space $\mathscr{K}(X)$! In fact, not only is integration a linear functional, it turns out that integration in general provides us with some of the "most important examples of a linear functional in mathematics," such as the Fourier coefficients of a periodic integrable function [FIS03].

This idea of integration being a linear functional is what motivates our study of the Riesz representation theorem. Since $\mathscr{K}(X)$ is, at its core, a vector space, and since vector spaces are generally well-understood, we know that there are many linear functionals that we can define on $\mathscr{K}(X)$ that may seem to have nothing to do with integration at all. This idea gives rise to many natural questions. Which linear functionals on $\mathscr{K}(X)$ can be represented in this way? All of them? Only some of them? It is clear that any regular Borel measure corresponds to one particular linear functional. But can one particular linear functional be represented by more than one regular Borel measure? It is these questions, and others like it, that led to the discovery of the Riesz representation theorem, which answers them all.

We will need the following concepts to state the Riesz representation theorem. Let X be a locally compact Hausdorff space. Firstly, we call a linear functional I on $\mathscr{K}(X)$ *positive* if, for every non-negative function $f \in \mathscr{K}(X)$, we have that $I(f) \geq 0$. Note that if I is a positive linear functional on $\mathscr{K}(X)$ and if f and g are functions in

$\mathscr{K}(X)$ so that $f \leq g$, then $g-f$ is non-negative, and so $I(g-f) = I(g)-I(f) \geq 0$, so that $I(f) \leq I(g)$. Thus, a positive linear functional is monotonic. Furthermore, if $f \in \mathscr{K}(X)$ is non-negative, then its integral is also non-negative by Theorem 3.2.9, so the mapping defined on $\mathscr{K}(X)$ by $f \mapsto \int f\,d\mu$ is actually a positive linear functional. Lastly, if U is an open subset of X, then we will write $f \prec U$ to denote that both $0 \leq f \leq \chi_U$ and $\operatorname{supp}(f) \subseteq U$.

We will now present one last fact which will be used in our proof of the main theorem.

Theorem 4.2.2. *Let X be a locally compact Hausdorff space, and let μ be a regular Borel measure defined on $\mathscr{B}(X)$. If U is an open subset of X, then*

$$\begin{aligned}\mu(U) &= \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } 0 \leq f \leq \chi_U\right\} \\ &= \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U\right\}.\end{aligned}$$

Proof. If $f \in \mathscr{K}(X)$ and $0 \leq f \leq \chi_U$, then, by Theorem 3.2.9, $\int f\,d\mu \leq \int \chi_U\,d\mu = \mu(U)$, so we must have that

$$\mu(U) \geq \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } 0 \leq f \leq \chi_U\right\}.$$

Also, if $f \in \mathscr{K}(X)$ and $f \prec U$, then necessarily $0 \leq f \leq \chi_U$, so

$$\begin{aligned}\mu(U) &\geq \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } 0 \leq f \leq \chi_U\right\} \\ &\geq \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U\right\}.\end{aligned}$$

Thus, if we prove that $\mu(U) \leq \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U\right\}$, then all three quantities must be equal and our theorem will be proved.

So, let α be an arbitrary real number so that $\alpha < \mu(U)$. Since μ is regular, we can use the inner regularity of μ to choose a compact subset K of U so that $\alpha < \mu(K)$. Thus, by Theorem 4.1.4, we can choose a function $f \in \mathscr{K}(X)$ so that $\chi_K \leq f \leq \chi_U$ and $\operatorname{supp}(f) \subseteq U$, and since $\chi_K \geq 0$, this means that $f \prec U$. Applying Theorem 3.2.9 to the inequality $\chi_K \leq f$ gives us that

$$\int f\,d\mu \geq \int \chi_K\,d\mu = \mu(K) > \alpha.$$

So, since $f \in \mathscr{K}(X)$, $f \prec U$, and $\alpha < \int f \, d\mu$, we must have that

$$\alpha < \sup \left\{ \int f \, d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U \right\}.$$

But α was an arbitrary real number less than $\mu(U)$, and so letting α approach $\mu(U)$ from below gives us that

$$\mu(U) \leq \sup \left\{ \int f \, d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U \right\}.$$

Therefore, our theorem has been proven. □

We now have all the machinery, language, and notation that we need to state and prove the main theorem of this project. This theorem answers all of the questions posed above regarding what kinds of linear functionals on $\mathscr{K}(X)$ can be represented by what kinds of regular Borel measures, stating that every single positive linear functional on $\mathscr{K}(X)$ is equivalent to integration with respect to a regular Borel measure, and that this measure is uniquely determined by the linear functional. This result is very useful, and is powerful enough that it is even possible to derive the Lebesgue measure λ on $\mathscr{B}(\mathbb{R})$ as a corollary [Rud87]. The proof of the theorem is long and complex, and will be presented in the next section.

Theorem 4.2.3 (The Riesz Representation Theorem for Linear Functionals). *Let X be a locally compact Hausdorff space, and let I be a positive linear functional on $\mathscr{K}(X)$. Then there is a unique regular Borel measure μ on X such that*

$$I(f) = \int f \, d\mu$$

holds for every f in $\mathscr{K}(X)$.

4.3 Proof of the Riesz Representation Theorem

In this section, X will always represent a locally compact Hausdorff space and I will represent a positive linear functional on $\mathscr{K}(X)$. We will split the proof into multiple parts in order to help with organization, and it will proceed as follows. First, we will show that, if there exists a regular Borel measure that satisfies the conditions in Theorem 4.2.3, then it must be unique. We will then construct a specific outer measure on X and prove that that its restriction to $\mathscr{B}(X)$ is a regular Borel measure on X. The last step will be to verify that our constructed measure does indeed satisfy the conclusion of Theorem 4.2.3. We have a big task ahead of us, so let us begin!

The regular Borel measure μ is unique

Let μ and ν be regular Borel measures on X such that $I(f) = \int f\,d\mu = \int f\,d\nu$ for every $f \in \mathscr{K}(X)$. If U is an open subset of X, Theorem 4.2.2 tells us that

$$\begin{aligned} \mu(U) &= \sup\left\{\int f\,d\mu : f \in \mathscr{K}(X) \text{ and } f \prec U\right\} \\ &= \sup\left\{\int f\,d\nu : f \in \mathscr{K}(X) \text{ and } f \prec U\right\} \\ &= \nu(U). \end{aligned}$$

Now, since μ and ν are regular, we can use their outer regularity and the fact that μ and ν agree on every open subset of X to get that, for all $A \subseteq \mathscr{B}(X)$,

$$\begin{aligned} \mu(A) &= \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\} \\ &= \inf\{\nu(U) : A \subseteq U \text{ and } U \text{ is open}\} \\ &= \nu(A). \end{aligned}$$

Thus, $\mu(A) = \nu(A)$ for every Borel subset A of X, meaning that they assign the same measure to every measurable subset of X. Therefore, $\mu = \nu$, so the regular Borel measure μ in the statement of Theorem 4.2.3 must be unique.

Constructing an outer measure on X

We will now define a function μ^* on the open subsets of X by letting, for all open sets $U \subseteq X$,

$$\mu^*(U) = \sup\{I(f) : f \in \mathscr{K}(X) \text{ and } f \prec U\}. \tag{4.1}$$

While at this point μ^* is neither a measure nor an outer measure, the notion of outer regularity suggests the next step, which will be to extend this function to every subset of X by defining, for any $A \subseteq X$,

$$\mu^*(A) = \inf\{\mu^*(U) : U \text{ is open and } A \subseteq U\}. \tag{4.2}$$

Before we proceed, we need to verify that (4.1) and (4.2) are consistent with each other in that they assign the same value to open sets. If we let V be an open subset of X, then (4.2) gives us that $\mu^*(V) = \inf\{\mu^*(U) : U \text{ is open and } V \subseteq U\}$. Now if $V \subseteq U$, then any function f in $\mathscr{K}(X)$ such that $f \prec V$ must also satisfy $f \prec U$, so that $\mu^*(V) \leq \mu^*(U)$, according to (4.1). Thus, since V is open and $V \subseteq V$, the set V is itself included in the sets U in (4.2), so we see that $\inf\{\mu^*(U) : U \text{ is open and } V \subseteq U\} = \mu^*(V)$, using (4.1). Therefore, (4.1) and (4.2) assign the same value to every open subset of X, and so our two definitions of μ^* are consistent with each other. Given this fact, for the rest of this proof we will choose whichever definition of μ^* will be more convenient, if we have the option to use either one. We will now prove that our function μ^* is actually an outer measure on X.

The function μ^* is an outer measure

To prove that μ^* is an outer measure on X, Definition 2.2.3 tells us that we need to show that $\mu^*(\emptyset) = 0$, that μ^* is monotonic, and that μ^* is countably subadditive.

Since $\emptyset$ is open in any topological space, we can apply (4.1) to see that

$$\mu^*(\emptyset) = \sup\{I(f) : f \in \mathscr{K}(X) \text{ and } f \prec \emptyset\}.$$

But if $f \in \mathscr{K}(X)$ and $f \prec \emptyset$, then by definition, $0 \leq f \leq \chi_\emptyset$ and $\operatorname{supp}(f) \subseteq \emptyset$, both of which force f to be the zero map 0, since $\chi_\emptyset = 0$ and $\overline{\{x \in X : f(x) \neq 0\}} \subseteq \emptyset$ implies that $f(x) = 0$ for every $x \in X$. Thus, we have that $\mu^*(\emptyset) = I(0)$. But I is linear, so for any $f \in \mathscr{K}(X)$,

$$I(0) = I(f - f) = I(f) - I(f) = 0.$$

Therefore, $\mu^*(\emptyset) = 0$.

Now let $A \subseteq B \subseteq X$. By (4.2), we have that

$$\mu^*(A) = \inf\{\mu^*(U) : U \text{ is open and } A \subseteq U\}$$

and that

$$\mu^*(B) = \inf\{\mu^*(U) : U \text{ is open and } B \subseteq U\}.$$

But since $A \subseteq B$, any open set containing B must also contain A, so that

$$\{\mu^*(U) : U \text{ is open and } B \subseteq U\} \subseteq \{\mu^*(U) : U \text{ is open and } A \subseteq U\}.$$

Therefore, since the infimum of any set is less than or equal to the infimum of any of its subsets, we have that $\mu^*(A) \leq \mu^*(B)$, and so μ^* is monotonic.

To prove that μ^* is countably subadditive, we need to show that, for any arbitrary sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of X, that

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

However, we will first show that μ^* is countably subadditive with respect to sequences of open subsets of X, and then use this fact to prove the case where we have an arbitrary sequence of subsets.

Let $\{U_n\}_{n=1}^{\infty}$ be a sequence of open subsets of X, and let f be a function in $\mathscr{K}(X)$ such that $f \prec \bigcup_{n=1}^{\infty} U_n$. By definition, this means that $0 \leq f \leq \chi_{\bigcup_{n=1}^{\infty} U_n}$ and that $\{U_n\}_{n=1}^{\infty}$ is an open cover of the set $\operatorname{supp}(f)$. Since $f \in \mathscr{K}(X)$, $\operatorname{supp}(f)$ is compact and so there must exist a finite subcover. So, let N be a positive integer such that $\operatorname{supp}(f) \subseteq \bigcup_{n=1}^{N} U_n$. We can now use Theorem 4.1.5 to choose functions $f_1, f_2, \ldots, f_N$ in $\mathscr{K}(X)$ so that $f = \sum_{n=1}^{N} f_n$ and so that $\operatorname{supp}(f_n) \subseteq U_n$ for all $n \in \{1, 2, \ldots, N\}$. Since $0 \leq f \leq \chi_{\bigcup_{n=1}^{\infty} U_n}$, we know that we can also choose each function f_n such that $0 \leq f_n \leq \chi_{U_n}$, so we have that $f_n \prec U_n$ for all $n \in \{1, 2, \ldots, N\}$. Since I is linear, this implies that

$$I(f) = I\left(\sum_{n=1}^{N} f_n\right) = \sum_{n=1}^{N} I(f_n).$$

Now, for all $n \in \{1, 2, \ldots, N\}$, $f_n \in \mathscr{K}(X)$ and $f_n \prec U_n$ imply that

$$I(f_n) \leq \sup\{I(g) : g \in \mathscr{K}(X) \text{ and } g \prec U_n\} = \mu^*(U_n),$$

by (4.1). Thus, we see that, since $\mu^*(U_n)$ is non-negative for all positive integers n,

$$I(f) = \sum_{n=1}^{N} I(f_n) \leq \sum_{n=1}^{N} \mu^*(U_n) \leq \sum_{n=1}^{\infty} \mu^*(U_n).$$

But since f was an arbitrary function in $\mathscr{K}(X)$ with $f \prec \bigcup_{n=1}^{\infty} U_n$, the above inequality must hold for every such f, so that

$$\sup\left\{I(f) : f \in \mathscr{K}(X) \text{ and } f \prec \bigcup_{n=1}^{\infty} U_n\right\} \leq \sum_{n=1}^{\infty} \mu^*(U_n).$$

But by (4.1),

$$\sup\left\{I(f) : f \in \mathscr{K}(X) \text{ and } f \prec \bigcup_{n=1}^{\infty} U_n\right\} = \mu^*\left(\bigcup_{n=1}^{\infty} U_n\right),$$

so we see that

$$\mu^*\left(\bigcup_{n=1}^{\infty} U_n\right) \leq \sum_{n=1}^{\infty} \mu^*(U_n).$$

Now let $\{A_n\}_{n=1}^{\infty}$ be an arbitrary sequence of subsets of X. We need to show that $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$. However, if $\sum_{n=1}^{\infty} \mu^*(A_n) = \infty$, then the required inequality is necessarily true. Thus, we will assume that $\sum_{n=1}^{\infty} \mu^*(A_n) < \infty$.

Let $\epsilon > 0$. Recall that for each integer $n \geq 1$, (4.2) gives that

$$\mu^*(A_n) = \inf\{\mu^*(U) : U \text{ is open and } A_n \subseteq U\}.$$

So, for each integer $n \geq 1$, we can choose an open set U_n such that $A_n \subseteq U_n$ (implying that $\mu^*(A_n) \leq \mu^*(U_n)$ since μ^* is monotonic) and such that

$$\mu^*(U_n) \leq \mu^*(A_n) + \frac{\epsilon}{2^n}.$$

Since μ^* is countably subadditive with respect to sequences of open sets, and by our choice of the sets U_n, we get that

$$\begin{aligned}
\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \mu^*\left(\bigcup_{n=1}^{\infty} U_n\right) \\
&\leq \sum_{n=1}^{\infty} \mu^*(U_n) \\
&\leq \sum_{n=1}^{\infty} \left(\mu^*(A_n) + \frac{\epsilon}{2^n}\right) \\
&= \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon.
\end{aligned}$$

Thus, since ϵ was arbitrary, we have that

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n),$$

so μ^* is countably subadditive. Since we also have that $\mu^*(\emptyset) = 0$ and μ^* is monotonic, μ^* is an outer measure by Definition 2.2.3. Our next step will be to prove that every set $A \in \mathscr{B}(X)$ is μ^*-measurable.

Every Borel subset of X is μ^*-measurable

First, let U be an open subset of X. By Definition 2.2.5, U is μ^*-measurable if and only if

$$\mu^*(A) = \mu^*(A \cap U) + \mu^*(A \cap U^c)$$

holds for every subset A of X. However, since $A \subseteq \big((A \cap U) \cup (A \cap U^c)\big)$ and μ^* is countably subadditive, then $\mu^*(A) \leq \mu^*(A \cap U) + \mu^*(A \cap U^c)$ must automatically be true. Thus, we need to prove that

$$\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c),$$

which is necessarily true if $\mu^*(A) = \infty$. So, let A be a subset of X such that $\mu^*(A) < \infty$, and let $\epsilon > 0$. By (4.2), $\mu^*(A) = \inf\{\mu^*(U) : U \text{ is open and } A \subseteq U\}$. Therefore, we can choose an open subset V of X such that $A \subseteq V$ (implying that $\mu^*(A) \leq \mu^*(V)$ since μ^* is monotonic) and such that $\mu^*(V) \leq \mu^*(A) + \epsilon$.

Now, since $V \cap U$ is open, we can use (4.1) to choose a function $f_1 \in \mathscr{K}(X)$ such that $f_1 \prec V \cap U$ and $I(f_1) \geq \mu^*(V \cap U) - \epsilon$. Let $K = \text{supp}(f_1)$. Since K is the closure of the set of points not sent to 0 by f_1, K is a closed set. So K^c is open, and so the set $V \cap K^c$ is open as well. Since $f_1 \prec V \cap U$, we have that $K \subseteq V \cap U$, so $K \subseteq U$ as well. Thus, $U^c \subseteq K^c$, so that $V \cap U^c \subseteq V \cap K^c$. Again by (4.1), we can choose a function $f_2 \in \mathscr{K}(X)$ so that $f_2 \prec V \cap K^c$ and $I(f_2) \geq \mu^*(V \cap K^c) - \epsilon$, and since μ^* is monotonic, we must have that $I(f_2) \geq \mu^*(V \cap U^c) - \epsilon$.

Recall that $A \subseteq V$. Now, by our choice of f_1 and f_2, we have that $f_1 + f_2 \prec V$, and so, since $f_1 + f_2 \subseteq \mathscr{K}(X)$, we can use (4.1), the linearity of I, and the monotonicity of μ^* to see that

$$\begin{aligned}
\mu^*(V) &\geq I(f_1 + f_2) \\
&= I(f_1) + I(f_2) \\
&\geq \mu^*(V \cap U) - \epsilon + \mu^*(V \cap U^c) - \epsilon \\
&= \mu^*(V \cap U) + \mu^*(V \cap U^c) - 2\epsilon \\
&\geq \mu^*(A \cap U) + \mu^*(A \cap U^c) - 2\epsilon.
\end{aligned}$$

But we chose V so that $\mu^*(V) \leq \mu^*(A) + \epsilon$, so we see that

$$\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c) - 3\epsilon.$$

Thus, since ϵ was arbitrary, $\mu^*(A) \geq \mu^*(A \cap U) + \mu^*(A \cap U^c)$. Therefore, every open subset U of X is μ^*-measurable. So, by Theorem 2.2.6, the collection $\mathscr{A}_{\mu^*}$ of μ^*-measurable sets is a σ-algebra that contains every open subset of X. However, by definition, $\mathscr{B}(X)$ is the *smallest* σ-algebra containing all of the open subsets of X, so $\mathscr{B}(X) \subseteq \mathscr{A}_{\mu^*}$. Therefore, every Borel subset A of X is μ^*-measurable.

Next, we will prove that restricting μ^* to the Borel σ-algebra on X yields a regular Borel measure, but before we can do this, we will need one more lemma.

Lemma 4.3.1. *Let X be a locally compact Hausdorff space, let I be a positive linear functional on $\mathscr{K}(X)$, and let μ^* be the outer measure defined by (4.1) and (4.2). Let A be a subset of X and let f be a function in $\mathscr{K}(X)$.*

1. *If $\chi_A \leq f$, then $\mu^*(A) \leq I(f)$.*

2. *If $0 \leq f \leq \chi_A$ and A is compact, then $I(f) \leq \mu^*(A)$.*

Proof. To prove part 1, let $\chi_A \leq f$, and let ϵ be a real number such that $0 < \epsilon < 1$. Now define the set U_ϵ by setting $U_\epsilon = \{x \in X : f(x) > 1 - \epsilon\}$. Since $f \in \mathscr{K}(X)$, f must be continuous, so U_ϵ is open as the preimage of the open set $(1-\epsilon, \infty)$ in $\mathbb{R}$. Let g be a function in $\mathscr{K}(X)$ so that $g \leq \chi_{U_\epsilon}$. By our choice of ϵ, we have that $0 < 1 - \epsilon < 1$.

If $x \in U_\epsilon$, then $g(x) \leq 1$ and $f(x) > 1 - \epsilon$, so $\frac{f(x)}{1-\epsilon} > 1$. Thus,

$$g(x) \leq \frac{1}{1-\epsilon} f(x).$$

If $x \notin U_\epsilon$, then $g(x) \leq 0$, and since $0 \leq \chi_A \leq f$ and $1 - \epsilon > 0$, we again have that

$$g(x) \leq \frac{1}{1-\epsilon} \chi_A(x) \leq \frac{1}{1-\epsilon} f(x).$$

Therefore, $g \leq \frac{1}{1-\epsilon} f$.

So, for each g in $\mathscr{K}(X)$ that satisfies $g \prec U_\epsilon$, we have that $g \leq \frac{1}{1-\epsilon} f$, so that, by (4.1) and the monotonicity and linearity of I,

$$\begin{aligned}
\mu^*(U_\epsilon) &= \sup\{I(g) : g \in \mathscr{K}(X) \text{ and } g \prec U_\epsilon\} \\
&\leq I\left(\frac{1}{1-\epsilon} f\right) \\
&= \frac{1}{1-\epsilon} I(f).
\end{aligned}$$

If $x \in A$, then $f(x) \geq \chi_A(x) = 1 > 1 - \epsilon$, so $x \in U_\epsilon$. Thus, $A \subseteq U_\epsilon$, so that $\mu^*(A) \leq \mu^*(U_\epsilon)$ by the monotonicity of μ^*. Combining this fact with the above inequality gives us that

$$\mu^*(A) \leq \frac{1}{1-\epsilon} I(f).$$

Therefore, since ϵ was arbitrary in the interval $(0,1)$, we see that $\frac{1}{1-\epsilon}$ is arbitrarily close to 1 from below, so that $\mu^*(A) \leq I(f)$, and part 1 of our lemma is proven.

Now let A be compact, $0 \leq f \leq \chi_A$, and let U be an open set such that $A \subseteq U$. Since $f \in \mathscr{K}(X)$, we have that $\mathrm{supp}(f)$ is closed and compact, and thus a subset of A. Therefore, $\mathrm{supp}(f) \subseteq U$, so $f \prec U$. So by (4.1),

$$I(f) \leq \sup\{I(g) : g \in \mathscr{K}(X) \text{ and } g \prec U\} = \mu^*(U).$$

But U was an arbitrary open set containing A, so $I(f) \leq \mu^*(U)$ for all such open sets. Thus, we must have that

$$I(f) \leq \inf\{\mu^*(V) : V \text{ is open and } A \subseteq V\},$$

so by (4.2), $I(f) \leq \mu^*(A)$, and part 2 of our lemma is proven. □

Now that we have this tool at our disposal, we can continue our proof. Recall that our next step is to prove that restricting μ^* to $\mathscr{B}(X)$ will result in a regular Borel measure on X.

The function $\mu^*|_{\mathscr{B}(X)}$ is a regular measure on X

Let μ be the restriction of μ^* to the Borel σ-algebra $\mathscr{B}(X)$, and let μ_1 be the restriction of μ^* to the collection $\mathscr{A}_{\mu^*}$ of μ^*-measurable sets. Then, by Theorem 2.2.6, μ_1 is automatically a measure on $\mathscr{A}_{\mu^*}$. We have already proven that every Borel subset of X is μ^*-measurable, that is, that $\mathscr{B}(X) \subseteq \mathscr{A}_{\mu^*}$, so μ and μ_1 must agree on $\mathscr{B}(X)$. Now, since μ_1 takes values in $[0,\infty]$, so does μ. Since $\mu_1(\emptyset) = 0$ and $\emptyset \in \mathscr{B}(X)$, we see that $\mu(\emptyset) = 0$. If $\{B_n\}_{n=1}^{\infty}$ is a disjoint sequence of subsets of $\mathscr{B}(X)$, then it is also a disjoint sequence of subsets of $\mathscr{A}_{\mu^*}$, and since μ_1 is a measure,

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} B_n\right) &= \mu_1\left(\bigcup_{n=1}^{\infty} B_n\right) \\
&= \sum_{n=1}^{\infty} \mu_1(B_n) \\
&= \sum_{n=1}^{\infty} \mu(B_n).
\end{aligned}$$

Therefore, by Definition 2.2.1, μ is a measure on $\mathscr{B}(X)$, that is, μ is a Borel measure on X.

Now we turn to the regularity of μ. Firstly, since μ is simply a restriction of μ^*, note that every result that we have already proven about μ^* applies to μ, and that (4.1) and (4.2) can be used as the (consistent) definitions of μ. Now, recall that, by Definition 4.1.7, μ is regular if and only if

1. for every compact subset K of X, $\mu(K) < \infty$,

2. for every set $A \in \mathscr{B}(X)$,

$$\mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\}, \text{ and}$$

3. for every open set $U \subseteq X$,

$$\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ and } K \text{ is compact}\}.$$

Let K be an arbitrary compact subset of X. By Theorem 4.1.4, there exists a function $f \in \mathscr{K}(X)$ so that $\chi_K \leq f$. Thus, by Part 1 of Lemma 4.3.1, $\mu(K) \leq I(f)$. But I is real-valued, so $\mu(K)$ is finite. Thus, μ is finite on every compact subset of X. (4.2) tells us exactly that $\mu(A) = \inf\{\mu(U) : A \subseteq U \text{ and } U \text{ is open}\}$ for every set $A \in \mathscr{B}(X)$, that is, that μ is outer regular.

Let U be an open subset of X and let f be a function in $\mathscr{K}(X)$ such that $f \prec U$. If $x \in \operatorname{supp}(f) \subseteq U$, then, since $0 \leq f \leq \chi_U$, we have that $0 \leq f(x) \leq 1 = \chi_{\operatorname{supp}(f)}(x)$. If $x \notin \operatorname{supp}(f) = \overline{\{x \in X : f(x) \neq 0\}}$, then $f(x) = 0 = \chi_{\operatorname{supp}(f)}(x)$. Thus, we have that $0 \leq f \leq \chi_{\operatorname{supp}(f)}$. Therefore, since $f \in \mathscr{K}(X)$ implies that $\operatorname{supp}(f)$ is compact, Part 2 of Theorem 4.3.1 tells us that $I(f) \leq \mu(\operatorname{supp}(f))$ for every $f \in \mathscr{K}(X)$ such that $f \prec U$.

So, we can use this fact along with (4.1) to get that

$$\begin{aligned}\mu(U) &= \sup\{I(f) : f \in \mathscr{K}(X) \text{ and } f \prec U\}\\ &\leq \sup\{\mu(\mathrm{supp}(f)) : f \in \mathscr{K}(X) \text{ and } f \prec U\}\\ &\leq \sup\{\mu(K) : K \text{ is compact and } K \subseteq U\},\end{aligned}$$

where the last inequality comes from the fact that if $f \in \mathscr{K}(X)$ with $f \prec U$, then $\mathrm{supp}(f)$ is a compact subset of U, and so $\mu(\mathrm{supp}(f))$ must be less than or equal to the supremum of μ applied to *all* compact subsets of U. To prove the reverse inequality, we notice that, for every compact set K with $K \subseteq U$, the monotonicity of μ gives that $\mu(U) \geq \mu(K)$, so that

$$\mu(U) \geq \sup\{\mu(K) : K \subseteq U \text{ and } K \text{ is compact}\}.$$

Therefore, $\mu(U) = \sup\{\mu(K) : K \subseteq U \text{ and } K \text{ is compact}\}$, so μ is inner regular.

Thus, since we have shown that μ is finite on compact sets and that μ is both outer regular and inner regular, μ is a regular Borel measure on X, by definition. The last step of our proof will be to verify that this measure is indeed the measure that satisfies the conditions of the Riesz representation theorem (Theorem 4.2.3).

$I(f) = \int f \, d\mu$ for every function $f \in \mathscr{K}(X)$

We start by letting f be a non-negative function in $\mathscr{K}(X)$, and letting $\epsilon > 0$. We now define a sequence $\{f_n\}_{n=1}^{\infty}$ of functions on X by letting, for each positive integer n and for each $x \in X$,

$$f_n(x) = \begin{cases} 0 & \text{if } f(x) \leq (n-1)\epsilon, \\ f(x) - (n-1)\epsilon & \text{if } (n-1)\epsilon < f(x) \leq n\epsilon, \\ \epsilon & \text{if } n\epsilon < f(x). \end{cases}$$

Let $x \in X$. Then $f(x)$ is a non-negative real number, since f is real-valued and non-negative by hypothesis. If $f(x) = 0$, then $f(x) \leq (n-1)\epsilon$ for every positive integer n, so that $\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} 0 = 0$, so $f(x) = \sum_{n=1}^{\infty} f_n(x)$. If $f(x) > 0$, then there exists a unique positive integer n_0 such that $(n_0 - 1)\epsilon < f(x) \leq n_0\epsilon$. Thus, for every positive integer n with $n < n_0$, we have that $n\epsilon < f(x)$, and for every positive integer n

with $n_0 < n$, we have that $f(x) \leq (n-1)\epsilon$. Thus, we see that

$$\begin{aligned}
\sum_{n=1}^{\infty} f_n(x) &= \sum_{n=1}^{n_0-1} \epsilon + \sum_{n=n_0}^{n_0} \left(f(x) - (n-1)\epsilon\right) + \sum_{n=n_0+1}^{\infty} 0 \\
&= (n_0 - 1)\epsilon + f(x) - (n_0 - 1)\epsilon + 0 \\
&= f(x).
\end{aligned}$$

Therefore, in all cases, $f = \sum_{n=1}^{\infty} f_n$.

For each positive integer n, the function f_n is clearly real-valued. Note that each of the functions that make up the piecewise defined function f_n are continuous, and, since $f(x) - (n-1)\epsilon$ equals 0 if $f(x) = (n-1)\epsilon$ and equals ϵ if $f(x) = n\epsilon$ (that is, the pieces of each function f_n coincide at their endpoints), each function f_n is continuous.

Now let n be a positive integer. By the definition of f_n and the non-negativity of f, we have that

$$\begin{aligned}
\operatorname{supp}(f_n) &= \overline{\{x \in X : f_n(x) \neq 0\}} \\
&= \overline{\{x \in X : (n-1)\epsilon < f(x) \leq n\epsilon \text{ or } n\epsilon < f(x)\}} \\
&= \overline{\{x \in X : f(x) > (n-1)\epsilon\}} \\
&\subseteq \overline{\{x \in X : f(x) > 0\}} \\
&= \overline{\{x \in X : f(x) \neq 0\}} \\
&= \operatorname{supp}(f).
\end{aligned}$$

But $f \in \mathscr{K}(X)$ means that $\operatorname{supp}(f)$ is compact. Thus, $\operatorname{supp}(f_n)$ is a closed subset of the compact set $\operatorname{supp}(f)$, and so $\operatorname{supp}(f_n)$ is also compact. Therefore, since each function f_n is real-valued, continuous, and has compact support, we have that, for every positive integer n, the function f_n is in $\mathscr{K}(X)$.

As was stated in the proof of Theorem 4.2.1, since $f \in \mathscr{K}(X)$, we have that f is a continuous function that is only non-zero on a compact subset of X, and recall that continuous functions defined on compact sets are bounded. Therefore, f is bounded, so let M be a real number so that $|f(x)| = f(x) \leq M$ for all $x \in X$ and let N be a positive integer so that $M \leq (N-1)\epsilon$. So if $n > N$, we have that $f(x) \leq (n-1)\epsilon$ for all $x \in X$ and so $f_n = 0$. Thus, $f = \sum_{n=1}^{N} f_n$.

Now, let $K_0 = \operatorname{supp}(f)$, and for every positive integer n, let $K_n = \{x \in X : f(x) \geq n\epsilon\}$. Let n be a positive integer, and let $x \in X$. If $x \in K_n$, then $f(x) \geq n\epsilon$, so

$f_n(x) = \epsilon$. Since $\epsilon\chi_{K_n}(x) = \epsilon$, we have that $\epsilon\chi_{K_n}(x) = f_n(x)$. If $x \notin K_n$, then $f(x) < n\epsilon$, so

$$f_n(x) = \begin{cases} 0 & \text{if } f(x) \leq (n-1)\epsilon, \\ f(x) - (n-1)\epsilon & \text{if } (n-1)\epsilon < f(x). \end{cases}$$

So regardless of the value of $f(x)$, we see that $f_n(x) \geq 0$. Since $\epsilon\chi_{K_n}(x) = 0$, we have that $\epsilon\chi_{K_n}(x) \leq f_n(x)$. Thus, in any case, $\epsilon\chi_{K_n} \leq f_n$.

We will now consider the functions f_n and $\epsilon\chi_{K_{n-1}}$. First, consider the case where $n = 1$. If $x \in K_{n-1} = K_0$, then $\epsilon\chi_{K_0}(x) = \epsilon$, and

$$f_n(x) = f_1(x) = \begin{cases} 0 & \text{if } f(x) = 0, \\ f(x) & \text{if } 0 < f(x) \leq \epsilon, \\ \epsilon & \text{if } \epsilon < f(x). \end{cases}$$

So regardless of the value of $f(x)$, we see that $f_1(x) \leq \epsilon$, and so that $f_1(x) \leq \epsilon\chi_{K_0}(x)$. If $x \notin K_0 = \text{supp}(f)$, then $f(x) = 0$, and so $f_1(x) = 0$. Also, $\epsilon\chi_{K_0}(x) = 0$, so $f_1(x) = \epsilon\chi_{K_0}(x)$. Therefore, $f_1 \leq \epsilon\chi_{K_0}$.

Now let $n \geq 2$. If $x \in K_{n-1}$, then $f(x) \geq (n-1)\epsilon$, and so

$$f_n(x) = \begin{cases} f(x) - (n-1)\epsilon & \text{if } f(x) \leq n\epsilon, \\ \epsilon & \text{if } f(x) > n\epsilon. \end{cases}$$

So regardless of the value of $f(x)$, we see that $f_n(x) \leq \epsilon$. Also, $\epsilon\chi_{K_{n-1}}(x) = \epsilon$, so we have that $f_n(x) \leq \epsilon\chi_{K_{n-1}}(x)$. If $x \notin K_{n-1}$, then $f(x) < (n-1)\epsilon$, so that $f_n(x) = 0$. Also, $\epsilon\chi_{K_{n-1}}(x) = 0$, so that $f_n(x) = \epsilon\chi_{K_{n-1}}(x)$. Thus, $f_n \leq \epsilon\chi_{K_{n-1}}$.

Therefore, $f_n \leq \epsilon\chi_{K_{n-1}}$ for all positive integers n. Combining this with the above result that $\epsilon\chi_{K_n} \leq f_n$ gives us that, for every positive integer n,

$$\epsilon\chi_{K_n} \leq f_n \leq \epsilon\chi_{K_{n-1}}.$$

Now, by applying Part 1 of Theorem 4.3.1 to the inequality $\epsilon\chi_{K_n} \leq f_n$, we see that $\epsilon\mu(K_n) \leq I(f_n)$ for every positive integer n. Note that K_0 is compact since $f \in \mathscr{K}(X)$ and, for every positive integer n, $K_n \subseteq \text{supp}(f)$ and K_n is closed, as the continuous preimage of the closed set $[n\epsilon, \infty)$ in $\mathbb{R}$. Thus, each set K_n is a closed subset of the compact set K_0, so each set K_n is compact. Therefore, by applying Part 2 of Theorem 4.3.1 to the inequality $f_n \leq \epsilon\chi_{K_{n-1}}$, we see that $I(f_n) \leq \epsilon\mu(K_{n-1})$ for every positive integer n. Thus, we have that, for every positive integer n,

$$\epsilon\mu(K_n) \leq I(f_n) \leq \epsilon\mu(K_{n-1}).$$

So, by letting n range from 1 through N, the linearity of I and the fact that $f = \sum_{n=1}^{N} f_n$ give us that

$$\begin{aligned}
\sum_{n=1}^{N} \epsilon\mu(K_n) &\leq \sum_{n=1}^{N} I(f_n) \\
&= I\left(\sum_{n=1}^{N} f_n\right) \\
&= I(f) \\
&\leq \sum_{n=1}^{N} \epsilon\mu(K_{n-1}) \\
&= \sum_{n=0}^{N-1} \epsilon\mu(K_n).
\end{aligned}$$

Now, using Theorem 3.2.9 to integrate each portion of the inequality $\epsilon\chi_{K_n} \leq f_n \leq \epsilon\chi_{K_{n-1}}$ (since integration is monotone) gives us that, for all positive integers n,

$$\int \epsilon\chi_{K_n}\, d\mu \leq \int f_n\, d\mu \leq \int \epsilon\chi_{K_{n-1}}\, d\mu.$$

By applying Theorem 3.2.9 again (since integration is linear), and by the definition of the integral of a characteristic function, we see that, for all positive integers n,

$$\epsilon\mu(K_n) \leq \int f_n\, d\mu \leq \epsilon\mu(K_{n-1}).$$

Again by letting n range from 1 through N and remembering that $f = \sum_{n=1}^{N} f_n$, and by using the linearity of the integral as provided by Theorem 3.2.9, we see that

$$\begin{aligned}
\sum_{n=1}^{N} \epsilon\mu(K_n) &\leq \sum_{n=1}^{N} \int f_n\, d\mu \\
&= \int \left(\sum_{n=1}^{N} f_n\right) d\mu \\
&= \int f\, d\mu \\
&\leq \sum_{n=1}^{N} \epsilon\mu(K_{n-1}) \\
&= \sum_{n=0}^{N-1} \epsilon\mu(K_n).
\end{aligned}$$

So, to summarize what we have so far, we have proven that

$$\sum_{n=1}^{N} \epsilon\mu(K_n) \leq I(f) \leq \sum_{n=0}^{N-1} \epsilon\mu(K_n)$$

and that

$$\sum_{n=1}^{N} \epsilon\mu(K_n) \leq \int f\,d\mu \leq \sum_{n=0}^{N-1} \epsilon\mu(K_n).$$

Thus, each of the real numbers $I(f)$ and $\int f\,d\mu$ lie in the real interval

$$\left[\sum_{n=1}^{N} \epsilon\mu(K_n), \sum_{n=0}^{N-1} \epsilon\mu(K_n)\right],$$

which has length

$$\begin{aligned}
\sum_{n=0}^{N-1} \epsilon\mu(K_n) - \sum_{n=1}^{N} \epsilon\mu(K_n) &= \epsilon\sum_{n=0}^{N-1} \mu(K_n) - \epsilon\sum_{n=1}^{N} \mu(K_n) \\
&= \epsilon\left(\sum_{n=0}^{N-1} \mu(K_n) - \sum_{n=1}^{N} \mu(K_n)\right) \\
&= \epsilon\left(\mu(K_0) - \mu(K_N)\right) \\
&\leq \epsilon\mu(K_0) \\
&= \epsilon\mu\left(\text{supp}(f)\right),
\end{aligned}$$

since $\mu(K_N)$ is positive. But, since $f \in \mathscr{K}(X)$, we have that supp(f) is compact, and since μ is regular, that $\mu(\text{supp}(f))$ must be finite. Thus, $I(f)$ and $\int f\,d\mu$ lie in an interval with length less than or equal to ϵM, for some fixed real number M, and since ϵ was an arbitrary positive number, the length of this interval must be arbitrarily close to 0. Therefore,

$$I(f) = \int f\,d\mu.$$

So, given that X is a locally compact Hausdorff space and that I is a positive linear functional on $\mathscr{K}(X)$, we have constructed a regular Borel measure on X so that $I(f) = \int f\,d\mu$ holds for every function f in $\mathscr{K}(X)$, and proved it is unique. Therefore, the proof of the Riesz representation theorem is complete. □

Chapter 5

Conclusion

The purpose of this thesis was to prove the Riesz representation theorem. To complete this goal, we had to build up the preliminary material and concepts that were required, not only to be able to present the lengthy and complex proof of the theorem, but also to understand the beautiful relationship between linear functionals and measures that the theorem highlights. After introducing some ideas and questions that were, historically, motivating factors for the investigation of measure theory, we began to work our way through the subject.

In Chapter 2, we began our study of measure theory by defining the fundamental concepts of a σ-algebra and a measure and familiarizing ourselves with their basic properties. By generalizing the idea of a measure to that of an outer measure, we were able to construct a particularly useful measure called Lebesgue measure, and investigated many of its interesting properties and applications, for instance, the surprising fact that, in a very meaningful sense, the size of the Cantor set, which is uncountable and closed, is zero. We also introduced the Borel σ-algebra, which was a concept that we returned to throughout this project.

In Chapter 3, we used the useful tool of measures to define integration in a very generalized way. After introducing the concept of measurable functions and working out some important results, our construction of the integral proceeded in three steps. Firstly, we defined simple functions and how to integrate them with respect to a measure. We then were able to extend this definition to any non-negative (and possibly infinite-valued) measurable function using facts concerning measurable functions. The last step of our construction was to extend the integral even further to arbitrary $[-\infty, \infty]$-valued measurable functions to establish the concept of Lebesgue integration. After investigating some useful and important properties of the Lebesgue integral, we turned to three particular theorems that provided conditions under which the integral can commute with pointwise-limits of sequences of measurable functions. After proving these powerful theorems, we were ready to move on to the Riesz representation theorem.

Chapter 4 was where we began our preparations for the theorem by first studying some concepts from topology, such as local compactness and the Hausdorff property, and then moving on to define the terminology and notation necessary for the statement of the theorem. After presenting a series of necessary theorems and lemmas, we were able to state the Riesz representation theorem, which says that, if X is a locally compact Hausdorff space, then for any positive linear functional I on the set of functions $\mathscr{K}(X)$, there exists a unique regular Borel measure μ on X so that the linear functional is essentially equivalent to the measure, in the sense that applying I to any function in $\mathscr{K}(X)$ yields the exact same result as integrating the function with respect to μ. We proved this remarkable and elegant theorem in a sequence of smaller steps. First, we showed that if such a measure does exist, then it must be unique. We then constructed a particular function μ^* on X, and proved that it was an outer measure. After proving that our outer measure gives a Borel measure when restricted to the Borel σ-algebra on X, we then proved that this measure is regular. The last step of the proof, and perhaps the most challenging, was to prove that our regular Borel measure satisfied the requirements of the Riesz representation theorem.

In conclusion, this project was a pleasure to complete, and it was exciting both to finally answer some long-standing questions of mine pertaining to analysis and to add some ingenious proof techniques to my own personal mathematical toolbox. Of course, the most valuable reward that I have gained from completing this thesis project is a deep understanding of and appreciation for the basics (and a few not-so-basics!) of measure theory and functional analysis. I have enjoyed every second of my studies so far, and I look forward to learning as much as I possibly can about measure theory and all other areas of mathematics as I continue my mathematical education and the life-long learning process that is part of being a mathematician.

Bibliography

[BN00] George Bachman and Lawrence Narici. *Functional Analysis.* Dover Publications, Mineola, New York, 2000.

[Coh97] Donald Cohn. *Measure Theory.* Birkhäuser, Boston, Massachusetts, 1997.

[FIS03] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra.* Prentice-Hall, Upper Saddle River, New Jersey, 2003.

[Gor94] Russel A. Gordon. *The Integrals of Lebesgue, Denjoy, Perron, and Henstock.* American Mathematical Society, Providence, Rhode Island, 1994.

[Mun06] James R. Munkres. *Topology.* Pearson Education, Upper Saddle River, New Jersey, 2006.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, New York, New York, 1976.

[Rud87] Walter Rudin. *Real and Complex Analysis, Third Edition.* Tata McGraw-Hill, New York, New York, 1987.