2006

From measure to integration

Sara Hernandez McLoughlin

From Measure to Integration

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

Sara Hernandez McLoughlin

December 2006

FROM MEASURE TO INTEGRATION

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

by

Sara Hernandez McLoughlin

December 2006

Approved by:

[redacted]

Belisario Ventura, Committee Chair

11/28/06

Date

[redacted]

Zahid Hasan, Committee Member

[redacted]

Rolland Trapp, Committee Member

[redacted]

Peter Williams, Chair,
Department of Mathematics

[redacted]

Joseph Chavez
Graduate Coordinator,
Department of Mathematics

Abstract

The notions of outer measure, Lebesgue measurable sets and Lebesgue measure are studied in detail. The existence of nonmeasurable sets is proven thus demonstrating that the family of Lebesgue measurable sets is properly contained in the power set of $\mathbb{R}$. Moreover, a complete description of the Cantor and generalized Cantor sets is given. The Cantor set along with the Cantor function are used to construct a measurable set that is not Borel; hence, showing that the class of Lebesgue measurable sets is larger than the class of Borel measurable sets. In addition, the generalized Cantor set is used to provide an example of an open set whose boundary has positive measure.

After developing Lebesgue integration over the real line, the Riemann integrable functions are classified as those functions whose set of points of discontinuity has measure zero. Then the convergence theorems are proven and it is shown how these theorems are valid under less stringent assumptions that are required for the Riemann integral. Finally, a detail analysis of abstract measure theory for general measure spaces is given.

Acknowledgements

First and foremost, I would like to thank Dr. Belisario Ventura for his advising on this project. I am grateful for all your help and patience. Thank you to my committee, Dr. Hasan and Dr. Trapp, for reading through this paper. Many thanks to all the professors I have had during my attendance at California State University San Bernardino.

I would also like to thank my husband Peter McLoughlin. Thank you for suggesting Lebesgue Measure and Integration as a topic for my thesis. I really enjoyed learning about it. Thank you for always taking time to answer all my questions.

Thank you to all my family and to my friends and colleagues at SAIL. Your encouragement and belief in me has made all the difference.

Table of Contents

Chapter 1

Introduction

Archimedes was the first to develop a theory of integration in the third century BC. His approach was to calculate the value of definite integrals by systematic methods. But this approach could only be applied to highly symmetric figures. Then Newton and Leibniz, in the sixteen hundreds, developed the method of antiderivatives; that is, for a function f on [a,b] and an antiderivative F of f such that $F' = f$ we have $\int_a^b f(x)dx = F(b) - F(a)$. The advantage of this method was that a large class of integrals could be calculated, many with great ease. The disadvantage was that it lacked a rigorous foundation. A rigorous theory of integration was developed by Cauchy in 1823. His definition of integration was similar to Riemann, but he argued that in order for the integral of a function on [a,b] to exist, f had to be continuous on [a,b]. Then in 1854, Riemann developed the well-known Riemann integral; for a bounded function f on [a,b], $\int_a^b f(x)dx = lim_{\|\triangle x\| \to 0} \sum_{i=1}^{n} f(\overline{x}_i)(x_i - x_{i-1})$, where $\overline{x}_i$ is an arbitrary point of $[x_{i-1}, x_i]$. Riemann concluded that the integral of a bounded function f on [a,b] exists as long as f is not too discontinuous [Bur98].

In 1902 Henry Lebesgue combined the notions of measure and integration. The result was a procedure for constructing the integral that was very different from Riemann's construction. Instead of partitioning the the domain of the function, he partitioned the range. Under Lebesgue theory of integration every Riemann integrable function is also Lebesgue integrable. However, the converse is not true. In the process of developing his theory of integration, Lebesgue developed the Lebesgue measure which is a generalization of the length of intervals to sets known as measurable sets [Bar66]. In chapter 2, a detailed analysis of the construction of the Lebesgue measure is given. Then in chapter 6 we use the measure theory developed in chapter 2 to define the Lebesgue integral. We also show in chapter 6 the classification of Riemann integrable functions as those functions whose sets of points of discontinuity have measure zero. Finally in chapter 8, we study general measure theory. We show how the procedure that Lebesgue used to create the Lebesgue measure can be done in general to create other measures.

Lebesgue integration turns out to be more powerful and has greater applications than Riemann integration. We will see in chapter 7 that one of the advantages of the Lebesgue integral over the Riemann integral lies in the facilitation of interchanging the limit and the integral. Lebesgue integration provided more general convergence theorems [Rud76]. The proofs of the Bounded Convergence, Monotone Convergence and Dominated Convergence theorems is given in chapter 7 along with four examples that demonstrate the differences between the two integrals.

We will also explore some interesting facts about measurable sets. One of which is the existence of nonmeasurable sets, given in chapter 4. Thus showing that the family of measurable sets is properly contained in the power set of $\mathbb{R}$. We also construct a measurable set that is not Borel in chapter 5. Moreover, in chapter 3, we study the Cantor and the generalized Cantor sets and we give an example of an open set whose boundary has positive measure.

Chapter 2

Lebesgue Measure

Lebesgue measure is a generalization of the length $\ell(I)$ of an interval I to more complex subsets of $\mathbb{R}$. We would like to define a set function m that assigns to each set E a nonnegative extended real number mE called the measure of E that has the following four properties:

1. mE is defined for each set E of real numbers
2. $mI = \ell(I)$, where I is an interval and $\ell(I)$ is its length.
3. If $\{E_n\}$ is a sequence of disjoint sets, then $m(\bigcup E_n) = \sum mE_n$
4. m is translation invariant, that is $m(y + E) = m(E)$ for any $y \in \mathbb{R}$

Property 3 above is usually referred to as countable additivity, and in order to have it satisfied, each set E should belong to a family M of subsets of $\mathbb{R}$, with M being a $\sigma - algebra$ of sets, definition 2.1 and 2.2 clarify this concept.

Definition 2.1. *A collection* $\boldsymbol{A}$ *of subsets of* X *is called an algebra of sets if (i)* $B \cup C$ *is in* $\boldsymbol{A}$ *whenever* B *and* C *are, and (ii)* B^c *is in* $\boldsymbol{A}$ *whenever* B *is.*

Definition 2.2. *An algebra* $\boldsymbol{A}$ *of sets is called a* $\sigma - algebra$*, if every union (and intersection) of a countable collection of sets in* $\boldsymbol{A}$ *is again in* $\boldsymbol{A}$.

2.1 Outer Measure

Definition 2.3. *We define the outer measure* m^*A *of a set* A *to be*
$m^*A = inf_{A \subseteq \bigcup I_n}(\sum \ell(I_n))$

where I_n is a countable collection of open intervals that cover A, and we are considering the sum of the length of the intervals in any such collection. The outer measure of A is the infimum of all such sums.

The outer measure m^* is a set function satisfying almost all 4 conditions we would like our desired set function to have. The almost part refers to the fact that the outer measure is not countably additive but rather countably subadditive that is $m^*(\cup A_n) \leq \sum m^* A_n$ [Roy88]. The next proposition lists the properties of the outer measure, the proof is postponed until chapter 8.

Proposition 2.4. *The outer measure has these properties:*

1. $m^*(E)$ *is defined for each set* E *of real numbers*

2. $m^*\emptyset = 0$

3. If $A \subset B$, *then* $m^*(A) \leq m^*(B)$

4. $m^*(I) = \ell(I)$

5. $m^*(\bigcup A_n) \leq \sum m^* A_n$

6. m^* *is translation invariant*

As we can see the advantage of the outer measure is that it is defined for all sets. The disadvantage is that it is not countably additive. It is actually impossible to construct a set function with all four properties mentioned in the introductory paragraph. Usually the first condition is weakened in order to retain the last three conditions, that is we have to reduce the family of sets on which the outer measure is defined in order to make it countably additive. To do this we use Caratheodory's definition given below.

Definition 2.5. *A set* E *is measurable if for each set* A *we have* $m^*A = m^*(A \cap E) + m^*(A \cap E^c)$, *where* E^c *represents the complement of* E.

Some sets that are known to be measurable are sets whose outer measure is zero. Hence we will prove the following lemma.

Lemma 2.6. *If* $m^*E=0$, *then* E *is measurable.*

Proof. By property 3 of the outer measure, $m^*(A \cap E) \leq m^*(E)$ since $A \cap E \subset E$. It follows $m^*(A \cap E) = 0$. Now $m^*(A \cap E^c) \leq m^*(A)$ since $A \cap E^c \subset A$. Moreover $m^*(A) = m^*((A \cap E) \cup (A \cap E^c)) \leq m^*(A \cap E) + m^*(A \cap E^c)$ by countable subadditivity.

Also $m^*(A) \leq m^*(A \cap E^c)$ since $m^*(A \cap E) = 0$. This implies $m^*(A) = m^*(A \cap E^c)$ since $m^*(A) \geq m^*(A \cap E^c)$ and $m^*(A) \leq m^*(A \cap E^c)$. Hence $m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$.
□

2.2 Lebesgue Measurable Sets

The sets obtained by using the Caratheodory's criteria form a σ-algebra which we will call M, and the outer measure restricted to this σ-algebra is countably additive. Moreover M contains the intervals. The proofs of these statements is given in chapter 8. In chapter 4 we will show that M is properly contained in the power set of $\mathbb{R}$; that is, there are subsets of $\mathbb{R}$ that are not elements of M. We call the elements of M Lebesgue measurable sets or just simply measurable sets[Bur98][Roy88].

Next we will take a look at a σ-algebra called the σ-algebra of Borel sets. The collection of Borel sets are due to Emile Borel who in 1898 came up with a measure on the Borel sets. We will see that the class of Borel measurable sets belongs to the class of Lebesgue measurable sets[Rud87]. We will show, in chapter 5, that the converse does not hold. The definition of Borel sets is given below, along with the proof that every Borel set is measurable. But first recall that a set O of real numbers is called open if for every x in O there is an open interval I such that $x \in I \subset O$. Moreover, a set F of real numbers is called closed if its complement is open.

Definition 2.7. *The collection B of Borel sets is the smallest $\sigma-$algebra which contains all the open sets. Moreover, it is also the smallest $\sigma-$ algebra which contains all of the closed sets and the smallest $\sigma-$ algebra that contains the open intervals.*

Theorem 2.8. *Every Borel set is measurable.*

Proof. M contains the family B of Borel sets since M is a $\sigma-$algebra containing the open intervals, and B is the smallest $\sigma-$algebra containing the open intervals. □

2.3 Lebesgue Measure

Definition 2.9. *If E is a measurable set, then the Lebesgue measure mE is defined to be the outer measure of E. That is, m is the restriction of m^* to the family M of measurable sets. We call this set function m, the Lebesgue measure.*

The Lebesgue measure is due to Henri Lebesgue(1902). Since m is the restriction of m^* to the family M of measurable sets, m is our sought after function, that is m is a set function defined on a family M of subsets of $\mathbb{R}$ containing the intervals, with M a $\sigma - algebra$, that satisfies the following conditions(proof given in chapter 8):

1. $mI = \ell(I)$
2. $m(\bigcup E_n) = \sum mE_n$ for each sequence $\{E_n\}$ of disjoint sets in M
3. m is translation invariant

From now on when measure is used we are referring to Lebesgue measure. Now that Lebesgue measure has been defined, let us find the measure of some sets. But before we do this, we give some definitions and examples.

Definition 2.10. *A set A is countable if there is a 1-1 mapping from A onto the set of natural numbers.*

Definition 2.11. *A set A is uncountable if A is neither finite nor countable.*

For example the rational numbers are a countable set while the irrational numbers are uncountable. The first sets we will measure are the countable sets which have measure zero. The proof follows.

Theorem 2.12. *Every countable set has measure zero. That is, if $A \subset \mathbb{R}$ is a countable set, then $m^*(A) = 0$.*

Proof. Since A is countable there exists an enumeration of A . Without loss of generality we may assume $A = \{x_n : n = 1, 2, \dots\}$. Let $\epsilon > 0$. Consider $C = \bigcup_{n=1}^{\infty} \left(x_n - \frac{\epsilon}{2^{n+1}}, x_n + \frac{\epsilon}{2^{n+1}}\right)$. By design $\{x_n\} \subset C$. Moreover, $\ell\left(x_n - \frac{\epsilon}{2^{n+1}}, x_n + \frac{\epsilon}{2^{n+1}}\right) = \frac{\epsilon}{2^n}$.
Hence $\sum_{n=1}^{\infty} \ell\left(x_n - \frac{\epsilon}{2^{n+1}}, x_n + \frac{\epsilon}{2^{n+1}}\right) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon$.
By definition $m^*(\{x_n\}) = inf\left\{\sum_{i=1}^{\infty} \ell(a_i, b_i) | \{x_n\} \subset \bigcup_{i=1}^{\infty}(a_i, b_i)\right\}$.
Therefore $m^*(\{x_n\}) \leq \epsilon$. Since $m^*(\{x_n\}) \geq 0$ and $\epsilon > 0$ and arbitrary, $m^*(\{x_n\}) = 0$. □

Note that every countable set is measurable since every countable set has measure zero and by lemma 2.6 every set of outer measure zero is measurable. The next sets we will measure are the Cantor and the generalized Cantor sets. We will do this in the next chapter. We will also show in the next chapter that the converse to theorem 2.12 is not true.

Chapter 3

The Cantor and the Generalized Cantor Sets

In this chapter we will study some special sets called the Cantor and the generalized Cantor sets and we will find their measures. Then in section 3.1 we will look at an application of the generalized Cantor sets. Let us begin by defining the Cantor and the generalized Cantor sets by means of a constructive process.

Definition 3.1. *The Generalized Cantor Set. Let* $0 < \alpha \leq 1$.

Step1: Divide [0,1], into two closed intervals of equal length by removing an open interval of length $\frac{\alpha}{3}$. *We are left with the closed set* $F_1(\alpha) = [0, d_{\alpha_1}] \cup [d_{\alpha_1} + \frac{\alpha}{3}, 1]$, *where* $2d_{\alpha_1} + \frac{\alpha}{3} = 1$.

Step 2: Remove an open interval of length $\frac{\alpha}{3^2}$ *from each closed interval of* $F_1(\alpha)$ *in such a way to obtain four closed intervals of equal length. We are left with the closed set* $F_2(\alpha) = [0, d_{\alpha_2}] \cup [d_{\alpha_2} + \frac{\alpha}{3^2}, d_{\alpha_1}] \cup [d_{\alpha_1} + \frac{\alpha}{3}, d_{\alpha_1} + d_{\alpha_2} + \frac{\alpha}{3}] \cup [d_{\alpha_1} + d_{\alpha_2} + \frac{\alpha}{3} + \frac{\alpha}{3^2}, 1]$, *where* $2d_{\alpha_2} + \frac{\alpha}{3^2} = d_{\alpha_1}$.

Step 3: Remove an open interval of length $\frac{\alpha}{3^3}$ *from each closed interval of* $F_2(\alpha)$ *in such a way to obtain* 2^3 *closed intervals of equal length.*

Continuing this way, at the nth step we remove an open interval of length $\frac{\alpha}{3^n}$ *from each of the closed intervals of* $F_{n-1}(\alpha)$ *in such a way to obtain* 2^n *closed intervals of equal lengths.*

The generalized Cantor set $C(\alpha)$ *is the intersection of the* $F_n(\alpha)'s$, *that is*

$C(\alpha) = \bigcap_{n=1}^{\infty} F_n(\alpha)$.

Note that for $\alpha = 1$, $C(1)$ is the Cantor set.

The following are some properties of the distances of the intervals removed during the construction of the generalized Cantor set. These properties will be used in the application section of this chapter. Let d_{α_n} = the length of each closed interval in $F_n(\alpha)$.

Proposition 3.2. $d_{\alpha_n} = \frac{1}{2}[d_{\alpha_{n-1}} - \frac{\alpha}{3^n}]$

Proof. By construction

$2d_{\alpha_1} + \frac{\alpha}{3} = 1$

$2d_{\alpha_2} + \frac{\alpha}{3^2} = d_{\alpha_1}$

$\vdots$

$2d_{\alpha_n} + \frac{\alpha}{3^n} = d_{\alpha_{n-1}}$

Solving for d_{α_n} gives $d_{\alpha_n} = \frac{1}{2}[d_{\alpha_{n-1}} - \frac{\alpha}{3^n}]$.□

Proposition 3.3. $d_{\alpha_n} < \frac{1}{2^{n-1}} d_{\alpha_1}$

Proof. For $n = 2$ we have, by Proposition 3.2, that $d_{\alpha_2} = \frac{1}{2}[d_{\alpha_1} - \frac{\alpha}{3}] = \frac{1}{2}d_{\alpha_1} - \frac{1}{2}\frac{1}{3} < \frac{1}{2}d_{\alpha_1}$. Assume, we have $d_{\alpha_n} < \frac{1}{2^{n-1}} d_{\alpha_1}$. Then by Proposition 3.2, $d_{\alpha_{n+1}} = \frac{1}{2}[d_{\alpha_n} - \frac{\alpha}{3^{n+1}}]$, so $d_{\alpha_{n+1}} = \frac{1}{2}d_{\alpha_n} - \frac{1}{2}\frac{\alpha}{3^{n+1}} < \frac{1}{2}d_{\alpha_n} < \frac{1}{2}\frac{1}{2^{n-1}}d_{\alpha_1} = \frac{1}{2^n}d_{\alpha_1}$. Therefore we must have $d_{\alpha_n} < \frac{1}{2^{n-1}}d_{\alpha_1}$. □

Proposition 3.4. $d_{\alpha_n} \geq \frac{\alpha}{3^n}$

Proof. When $n = 1$, we have $2d_{\alpha_1} + \frac{\alpha}{3} = 1$. Suppose $d_{\alpha_1} < \frac{\alpha}{3}$ then $2d_{\alpha_1} + \frac{\alpha}{3} < \frac{2\alpha}{3} + \frac{\alpha}{3} = \alpha \leq 1$ Contradiction. Hence $d_{\alpha_1} \geq \frac{\alpha}{3}$.

Assume is true for n-1, then $d_{\alpha_{n-1}} \geq \frac{\alpha}{3^{n-1}}$ and $2d_{\alpha_n} + \frac{\alpha}{3^n} = d_{\alpha_{n-1}}$. Suppose $d_{\alpha_n} < \frac{\alpha}{3^n}$ then $d_{\alpha_{n-1}} = 2d_{\alpha_n} + \frac{\alpha}{3^n} < \frac{2\alpha}{3^n} + \frac{\alpha}{3^n} = \frac{3\alpha}{3^n} = \frac{\alpha}{3^{n-1}} \leq d_{\alpha_{n-1}}$. Hence $d_{\alpha_{n-1}} = 2d_{\alpha_n} + \frac{\alpha}{3^n} < d_{\alpha_{n-1}}$. Contradiction. Hence $d_{\alpha_n} \geq \frac{\alpha}{3^n}$. □

Proposition 3.5. $d_{\alpha_n} > 0$ *and* $\lim d_{\alpha_n} = 0$

Proof. By previous propositions, we have $\frac{\alpha}{3^n} \leq d_{\alpha_n} \leq \frac{1}{2^n}d_{\alpha_1}$.

Since $d_{\alpha_1} < 1$, we have $\frac{\alpha}{3^n} \leq d_{\alpha_n} \leq \frac{1}{2^n}d_{\alpha_1} < \frac{1}{2^n}$ which implies $\frac{\alpha}{3^n} \leq d_{\alpha_n} < \frac{1}{2^n}$.

Hence, $\lim_{n\to 0} \frac{\alpha}{3^n} \leq \lim_{n\to 0} d_{\alpha_n} < \lim_{n\to 0} \frac{1}{2^n}$ which gives $\lim_{n\to 0} d_{\alpha_n} = 0$.□

Proposition 3.6. $d_{\alpha_n} < d_{\alpha_{n-1}}$

Proof. $d_{\alpha_n} = \frac{1}{2}[d_{\alpha_{n-1}} - \frac{\alpha}{3^n}]$ which implies $d_{\alpha_n} < \frac{1}{2}d_{\alpha_{n-1}}$. Hence $d_{\alpha_n} < d_{\alpha_{n-1}}$.□

Now we are ready to find the measures of the generalized Cantor sets.

Theorem 3.7. $m^*\big(C(\alpha)\big) = 1 - \alpha$.

Proof. Let $D_n(\alpha)$ be the union of the open intervals removed at the nth step, and let $D(\alpha) = \bigcup_{n=1}^{\infty} D_n(\alpha)$. Now $[0,1] = C(\alpha) \cup D(\alpha)$ with $C(\alpha) \cap D(\alpha) = 0$. Since $D(\alpha)$ is a countable union of open sets, it is measurable. Now the complement of the set $D(\alpha)$ intersected with [0,1] is $C(\alpha)$. Since $C(\alpha)$ is the complement of a measurable set, it is measurable. Hence $m^*\big([0,1]\big) = m^*\big(C(\alpha)\big) + m^*\big(D(\alpha)\big)$.

Now the sum of the lengths of the intervals removed is the geometric series $\frac{\alpha}{3} + 2\frac{\alpha}{3^2} + 2^2\frac{\alpha}{3^3} + \cdots + 2^{n-1}\frac{\alpha}{3^n} + \cdots = \sum_{n=0}^{\infty} \alpha\frac{2^n}{3^{n+1}} = \frac{\alpha}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n = \frac{\frac{\alpha}{3}}{1-\frac{2}{3}} = \frac{\alpha}{3}3 = \alpha$.
So $m^*\big(D(\alpha)\big) = \alpha$ and since $m^*\big([0,1]\big) = 1$, we have $1 = m^*\big(C(\alpha)\big) + \alpha$.
Hence $m^*\big(C(\alpha)\big) = 1 - \alpha$. □

As a consequence of the theorem we have that the Cantor set has measure zero since when $\alpha = 1$, $m^*\big(C(1)\big) = 1 - 1 = 0$. Now the Cantor set is an uncountable set, the proof of this will be postponed until chapter 5. Hence we have found an uncountable set with measure zero!
From now on let us refer to $C(1)$ as just C, $D_n(1)$ as D_n, and $D(1)$ as D.

3.1 An Example of an Open Set whose Boundary has Positive Measure

In this section we will use the generalized Cantor sets to help us answer the following question: Does there exist an open set whose boundary has positive measure?

Let's explore this problem in detail.

Definition 3.8. *The closure of a set G, $\overline{G}$, is the intersection of all closed sets containing G.*

Definition 3.9. *The boundary of an open set O is $Bd(O) = \overline{O} - O$*

First note that any open set can be written as $O = \bigcup_{i=1}^{\infty}(a_i, b_i)$ where $(a_i, b_i) \bigcap (a_j, b_j) = \emptyset$ when $i \neq j$. Let $O_n = \bigcup_{i=1}^{n}(a_i, b_i)$ and note $\overline{O_n} = \{[a_i, b_i] \mid 1 \leq i \leq n\}$, and $Bd(O_n) = \{a_i, b_i \mid 1 \leq i \leq n\}$. It therefore seems reasonable to predict $Bd(O) = \bigcup_{n=1}^{\infty} Bd(O_n)$, and $m(Bd(O)) = 0$ since $\bigcup_{n=1}^{\infty} Bd(O_n)$ is countable. However, we well show that this reasoning is faulty.

The problem is that for a finite number of disjoint open intervals the boundary is precisely the endpoints of the intervals but when we are dealing with a countable number of disjoint open intervals this may not be true which means that the boundary is made up of more than just endpoints.

There exist an open set whose boundary has positive measure and that open set is $D(\alpha)$ with $\alpha \neq 1$; that is, the open set obtained during the Cantor process. It follows $D(\alpha) = \bigcup_{n=1}^{\infty} D_n(\alpha)$ where $D_n(\alpha)$ is the union of the open intervals removed at the nth step in the construction of the generalized Cantor set. The proof of this is given in theorem 3.12, but to prove theorem 3.12 we need lemma 3.10 and corollary 3.11.

Lemma 3.10. $\overline{D(\alpha)} = [0, 1]$

Proof. Since $a \in \overline{D(\alpha)}$ if and only if there exist a $x_n \subset D_\alpha$ such that $lim\, x_n = a$. It suffices to show that for any $a \in C(\alpha)$ there exists a sequence $\{x_n\}$ in $D(\alpha)$ such that $lim\, x_n = a$. Since $D(\alpha) = \bigcup_{n=1}^{\infty} D_n(\alpha)$, let $D_n(\alpha) = \bigcup_{i=1}^{2^{n-1}}(a_i, b_i)$ and $k_{in} = \frac{a_i+b_i}{2}$. Let $a \in C(\alpha)$. There exist $1 \leq m \leq 2^{n-1}$ such that $d(k_{mn}, a) \leq d(k_{im}, a)$ for $1 \leq i \leq 2^{n-1}$. Define $x_n = k_{mn}$ then $\{x_n\} \subset D(\alpha)$.

By proposition 3.5, $lim\, d_{\alpha_n} = 0$ and $lim\, \frac{\alpha}{3^n} = 0$. Hence for any $\epsilon > 0$ there exists an N such that $d_{\alpha_N} + \frac{\alpha}{3^N} < \epsilon$. Since $a \in C(\alpha)$, "a" is in precisely one closed interval of $F_N(\alpha)$. Moreover the length of this closed interval is d_{α_N}. In $D_N(\alpha)$ each of the open intervals has length $\frac{\alpha}{3^N}$ and there is at least one open interval directly to the right or left of the closed interval containing a. It follows $d(x_N, a) < d_{\alpha_N} + \frac{\alpha}{3^N}$. Hence for any $n > N$, $d_{\alpha_n} < d_{\alpha_N}$, $\frac{\alpha}{3^n} < \frac{\alpha}{3^N}$ and we have $d(x_n, a) < d_{\alpha_n} + \frac{\alpha}{3^n} < d_{\alpha_N} + \frac{\alpha}{3^N} < \epsilon$. It follows $lim\, x_n = a$. □

Corollary 3.11. $Bd(D(\alpha)) = C(\alpha)$

Proof. $Bd(D(\alpha)) = \overline{D(\alpha)} - D(\alpha)$.

Now $Bd(D(\alpha)) = [0,1] - D(\alpha)$ since $\overline{D(\alpha)} = [0,1]$.
Hence $Bd(D(\alpha)) = C(\alpha)$. □

Theorem 3.12. *There exist an open set whose boundary has positive measure.*

Proof. Let $D(\alpha)$ be the open set obtained in the construction of the generalized Cantor set with $\alpha \neq 1$. By corollary 3.11, we have that $Bd(D(\alpha)) = C(\alpha)$. Since $m^*(C(\alpha)) = 1 - \alpha$, we have $m^*\big(Bd(D(\alpha))\big) = 1 - \alpha$. □

As we have seen, the Cantor and the generalized Cantor sets are very important sets for they provide counterexamples to different propositions we might believe are true. We will see in chapter 5 yet another application of the Cantor set, and the generalized Cantor sets. This time they help us find a measurable set that is not Borel. But in order to find this measurable set that is not Borel, we fist need to prove the existence of a nonmeasurable set.

Chapter 4

Nonmeasurable Sets

In 1905 Vitali was the first mathematician to discover Lebesgue nonmeasurable sets, hence showing that M is properly contained in the power set of $\mathbb{R}$. The construction of every Lebesgue nonmeasurable set of real numbers requires the use of the Axiom of Choice which states that for any nonempty collection C of sets there is a choice function f such that $f(A) \in A$ for each $A \in C$. That it is actually impossible to construct a Lebesgue nonmeasurable set without the Axiom of Choice was proved by Solovay in 1970 [Bur98][Roy88].

The goal of this chapter is to prove that every set of positive measure contains a nonmeasurable set. Before doing the general proof we will first prove the case when the set of positive measure is the interval [0,1], this will be done in theorem 4.1 and corollary 4.2 which say that if we can break the interval [0,1] into a countable union of disjoint sets with the same outer measure then one of those sets has to be nonmeasurable. Theorem 4.1 and corollary 4.2 provide a guide for the construction of the general nonmeasurable set.

Note that by $\bigcup E_i^{disj}$ we mean the sets E_i's are disjoint.

Theorem 4.1. *If* $[0,1] = \bigcup E_i^{disj}$ *with* E_i *measurable, then* $m(E_i) \neq m(E_j)$ *for some* $i \neq j$.

Proof. Assume $m(E_i) = m(E_j)$ for all i and j, then either $m(E_i) = 0$ or $m(E_i) > 0$.
Case 1: $m(E_i) = 0$ then $m(\bigcup E_i) = 1 \neq \sum m(E_i) = 0$. Contradiction.
Case 2: $m(E_i > 0$ then $m(\bigcup E_i) = 1 \neq \sum m(E_i) = \infty$. Contradiction.

Hence, $m(E_i) \neq m(E_j)$ for some $i \neq j$. □

Corollary 4.2. *If* $[0,1] = \bigcup A_i^{disj}$ *such that* $m^*(A_i) = m^*(A_j)$ *for all* i, j *then there exists an* i *such that* A_i *is a nonmeasurable set.*

Proof. Assume A_i is measurable for all i, then by theorem 4.1 $m(A_i) \neq m(A_j)$ for some $i \neq j$. Moreover, $m^*(A_i) = m(A_i)$ and $m^*(A_j) = m(A_j)$. Hence, $m^*(A_i) \neq m^*(A_j)$. Contradiction. Hence, there exists an i such that A_i is a nonmeasurable set.□

Theorem 4.3. *If* E *is measurable and* $m(E) \neq 0$ *then there exist* $E_1 \subset E$ *such that* E_1 *is measurable,* $0 < m(E_1) \leq 1$ *and* $E_1 \subset (n, n+1)$ *for some* $n \in Z$.

Proof. $\mathbb{R} = \bigcup_{n\in Z}(n, n+1) \cup Z$
$E = E \cap \mathbb{R}$
$E = E \cap \left(\bigcup_{n\in Z}(n, n+1) \cup Z\right)$
$E = \left(\bigcup_{n\in Z}(n, n+1) \cap E\right) \cup (Z \cap E)$
$m(E) = m\left(\bigcup_{n\in Z}(n, n+1) \cap E\right) + m(Z \cap E)$
$m(E) = \sum_{n\in Z} m\left((n, n+1) \cap E\right)$
Since $m(E) > 0$, there exist n such that $m\left((n, n+1) \cap E\right) > 0$. □

And now we are ready to prove the general case.

Theorem 4.4. *If* E *is a measurable set with* $m(E) \neq 0$, *then* E *contains a nonmeasurable set.*

Proof. Using theorem 4.3, we can assume that there exist $E_1 \subset E$ such that E_1 is measurable, $0 < m(E_1) \leq 1$ and $E_1 \subset (n, n+1)$ for some $n \in Z$. It suffices to show E_1 contains a nonmeasurable set.

Define the equivalence relation on E_1 by $x \sim y$ if and only if $x - y \in Q \cap (-1, 1)$. Form the equivalent classes as follows: $\overline{X} = \{x + \frac{a}{b} \mid \frac{a}{b} \in Q\} \cap E_1$. Notice that each equivalent class will have a countable number of elements since it is indexed by $Q \cap (-1, 1)$. Also the collection of distinct equivalent classes is uncountable since E_1 is uncountable.

Let P be the set whose elements consist of exactly one element from each equivalent class, by the axiom of choice such a set exists. Let $\{r_i\} = Q \cap (-1, 1)$. Form the sets $P + r_i = \{x + r_i \mid x \in P \text{ and } I \in N\}$. Notice that by design $E_1 = \{\bigcup_{i=1}^{\infty} P + r_i\} \cap E_1$.

Lemma 1: $(P + r_i) \cap (P + r_j) = \emptyset$, where $i \neq j$.

Proof of lemma 1: Assume $(P + r_i) \cap (P + r_j) \neq \emptyset$. Then for some x_1 and $x_2 \in P$, we have $x_1 + r_i = x_2 + r_j$. This implies $x_1 - x_2$ is a rational number. Hence x_1 and x_2 are in the same equivalent classes. However by construction of P, x_1 and x_2 are in different equivalent classes. Hence $(P + r_i) \cap (P + r_j) = \emptyset$.

Lemma 2: $\bigcup_{i=1}^{\infty} P + r_i \subset (n-1, n+2)$.

Proof of lemma 2: For any $x \in P$ and r_i, we have $n < x < n+1$ and $-1 < r_i < 1$. Hence $n - 1 < x + r_i < n + 2$.

If P is measurable then so is $P + r_i$ and $m(P) = m(P + r_i)$ since measurable sets are translation invariant. Moreover, $\bigcup_{i=1}^{\infty} P + r_i$ is measurable since measurable sets form a $\sigma - algebra$. Now $m\left(\bigcup_{i=1}^{\infty} P + r_i\right) = \sum_{i=1}^{\infty} m(P + r_i) = \sum_{i=1}^{\infty} m(P)$ since $P + r_i \cap P + r_j = \emptyset$, for $i \neq j$.

Now $E_1 \subset \bigcup_{i=1}^{\infty} P + r_i \subset (n-1, n+2)$.

Suppose P is measurable then $m(E_1) \leq m\left(\bigcup_{i=1}^{\infty} P + r_i\right) \leq m\big((n-1, n+2)\big)$ implies $m(E_1) \leq \sum_{i=1}^{\infty} m(P + r_i) \leq m\big((n-1, n+2)\big)$ and $m(E_1) \leq \sum_{i=1}^{\infty} m(P) \leq m\big((n-1, n+2)\big)$ with $0 < \sum_{i=1}^{\infty} m(P) \leq 3$. If $\sum_{i=1}^{\infty} m(P) \leq 3$ then $m(P) = 0$, and if $0 < \sum_{i=1}^{\infty} m(P)$, then $m(P) > 0$. Contradiction. Hence P is a nonmeasurable set. □

Corollary 4.5. *Any set of positive outer measure contains a nonmeasurable set.*

Proof. Let E be the set of positive outer measure.

Case 1: If E is nonmeasurable, then done.

Case 2: If E is measurable then by theorem 4.3, there exist $E_1 \subset E$ such that E_1 is measurable, $0 < m(E_1) \leq 1$ and $E_1 \in (n, n+1)$ for some $n \in Z$. By theorem 4.4, E_1 contains a nonmeasurable set. Hence E contains a nonmeasurable set. □

Chapter 5

A Measurable Set that is not Borel

In chapter 2, we learned that every Borel set is measurable and that the collection of Borel sets is the smallest sigma algebra that contains the open sets. It seems as if the collection of Borel sets account for all conceivable sets, thus all possible measurable sets. For some years it wasn't clear if there existed Lebesgue measurable sets that were not Borel measurable, but in 1914 this question was put to rest by Suslin, a Russian mathematician. Suslin proved the existence of a Lebesgue measurable set that is not Borel by constructing such a set [Bur23][Rud76]. Hence we have that the Borel measurable sets are a proper subset of Lebesgue measurable sets!

In order to show that there exists a measurable set which is not Borel, it will be sufficient to show the following two things: First, if B is a Borel set and g a measurable function then $g^{-1}(B)$ is measurable; and second, there exists a measurable bijection g and measure zero set A such that $g^{-1}(A)$ is a measurable set with $m(g^{-1}(A)) > 0$. Using these two results and the fact that every set of positive outer measure contains a nonmeasurable set we will be able to construct a measurable set that is not Borel. We will see that the measurable function we need uses the Cantor ternary function which will be explained in section 2. We will also see how the Cantor set plays a key role in this proof. Hence sections one through three are devoted to laying the foundations. Finally in section 4, we will put it all together.

5.1 Measurable Functions

Definition 5.1. *An extended real-valued function f, defined on a Lebesgue measurable set of real numbers E, is said to be Lebesgue measurable on E if $f^{-1}((c,\infty]) = \{x : f(x) > c\}$ is a Lebesgue measurable subset of E for every real number c.*

Proposition 5.2. *Let f be an extended real-valued function whose domain is a Lebesgue measurable set of real numbers E, and c is any real number. Then the following statements are equivalent:*

1. f is a Lebesgue measurable function on E.
2. $f^{-1}((c,\infty]) = \{x : f(x) > c\}$ is a Lebesgue measurable subset of E.
3. $f^{-1}((c,\infty]) = \{x : f(x) \geq c\}$ is a Lebesgue measurable subset of E.
4. $f^{-1}([-\infty,c)) = \{x : f(x) < c\}$ is a Lebesgue measurable subset of E.
5. $f^{-1}([-\infty,c)) = \{x : f(x) \leq c\}$ is a Lebesgue measurable subset of E.

Some examples of measurable functions are continuous functions. The proof follows.

Theorem 5.3. *If f is a continuous function, then f is measurable.*

Proof. Recall, f is continuous if and only if $f^{-1}(O)$ is open for all open sets O. Since open intervals are open sets, we have f^{-1} of open intervals is open for all open intervals. Hence since open sets are measurable, f is a measurable function.□

The following is a relationship between measurable functions and Borel sets that is key in proving the existence of a measurable set that is not Borel.

Theorem 5.4. *Let f be measurable and B a Borel set. Then $f^{-1}(B)$ is measurable.*

Proof. Let $f : E \mapsto \mathbb{R}$, where E is a measurable set. Let $S = \{A \subset \mathbb{R} \mid f^{-1}(A)$ is measurable $\}$. Note if $A_i \in S$ then the following are true:
1) $f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(A_i)$which is measurable.
2) $f^{-1}\left(\bigcap_{i=1}^{\infty} A_i\right) = \bigcap_{i=1}^{\infty} f^{-1}(A_i)$ which is measurable.
3) $f^{-1}(\mathbb{R} - A) = f^{-1}(\mathbb{R}) - f^{-1}(A_i)$ which is measurable.

Also for $a < b$ we have $f^{-1}((a,\infty)) = \{x \in E | f(x) > a\}$ and $f^{-1}((-\infty,b)) = \{x \in E | f(x) < b\}$ are measurable. Hence, $f^{-1}(a,\infty) \cap f^{-1}(-\infty,b) = f^{-1}((a,\infty) \cap$

$(-\infty, b)) = f^{-1}(a,b)$ is measurable. Since the Borel sets are the smallest σ-algebra which contain open intervals, we must have $B \subset S$ where B is the collection of Borel sets. □

Finally we will show that the composition of a measurable function with a continuous function is measurable. But first we need this lemma.

Lemma 5.5. $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$

Proof. $x \in (g \circ f)^{-1}(A)$
$\iff (g \circ f)(x) \in A$
$\iff g(f(x)) \in A$
$\iff f(x) \in g^{-1}(A)$
$\iff x \in f^{-1}(g^{-1}(A))$
Hence $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. □

Theorem 5.6. *If f is a measurable real-valued function and g is a continuous function defined on $(-\infty, \infty)$, then $g \circ f$ is measurable.*

Proof. Let A be an open interval. By lemma 5.5, we have $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. Since g is a continuous function, $g^{-1}(A)$ is an open set. Now since an open set is a Borel set and f is measurable, it follows from theorem 5.3 that f^{-1} of a Borel set is measurable. Hence $g \circ f$ is measurable. □

5.2 The Ternary Representation of the Cantor Numbers

In chapter 3 we learned that the Cantor numbers are obtained by intersecting all the closed intervals removed during the Cantor process, i.e. $C = \bigcap_{n=1}^{\infty} F_n$. In this section we will study another way of representing the Cantor numbers, namely its ternary representation. Let us first take a look at what is meant by the binary expansion and ternary expansion of a real number $x \in (0,1)$.

Definition 5.7. *Let p be an integer greater than 1, and x a real number $0 < x < 1$. There is a sequence (a_n) of integers with $0 \leq a_n \leq p$ such that $x = \sum_1^{\infty} \frac{a_n}{p^n}$ and this sequence is unique except when x is of the form $\frac{q}{p^n}$, in which case there are exactly two such sequences. Every number in the interval $(0,1)$ has at least one, and at most two,* *expansions. If $p = 2$, this sequence is called the binary expansion of x. For $p = 3$ it is called the ternary expansion.*

The next two theorems classify all the numbers with two ternary expansions.

Theorem 5.8. *$x \in [0,1]$ has two ternary expansions if and only if $x = \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{k}{3^m}$ where $k = 1$ or $k = 2$ and $a_n \in \{0, 1, 2\}$.*

Proof. Suppose $x = \sum_{n=1}^{\infty} \frac{b_n}{3^n} = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$ where a_n, $b_n \in \{0, 1, 2\}$ and $\{a_n\} \neq \{b_n\}$. Let M be the smallest integer such that $b_m \neq a_m$. Without loss of generality we may assume $a_m > b_m$. Now $0 = \sum_{n=1}^{\infty} \frac{a_n - b_n}{3^n} \geq \frac{1}{3^m} + \sum_{n=m+1}^{\infty} \frac{a_n - b_n}{3^n} \geq \frac{1}{3^m} - [\sum_{n=0}^{\infty} \frac{2}{3^{m+1+n}}] = 0$. Hence, we must have $b_{n+m} = 2$ and $a_{n+m} = 0$ for $n \geq 1$. Moreover, $a_m = 1$ or $a_m = 2$.

Conversely, suppose $x \in \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{k}{3^m}$ where $k = 1$ or 2. Let $b_n = a_n$ for $n \leq m - 1$, $b_m = 0$ if $k = 1$ or $b_m = 1$ if $k = 2$ and $b_{n+m} = 2$ for $n \geq 1$. By design $x = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$. □

Theorem 5.9. *Let a_n, $b_n \in \{0, 1, 2\}$ for all n. If $x = \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{k}{3^m}$ where $k = 1$ or 2, and $b_n \neq a_n$ for some $n < k$ then $\sum_{n=1}^{\infty} \frac{b_n}{3^n} \neq x$.*

Proof. Suppose $x = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$. Let r be the smallest integer such that $b_n \neq a_n$.

Case 1: $b_r > a_r$. Then $0 = \sum_{n=1}^{\infty} \frac{b_n - a_n}{3^n} = \frac{b_r - a_r}{3^r} + \sum_{n=r+1}^{m-1} \frac{b_n - a_n}{3^n} + \frac{b_m - a_m}{3^m} + \sum_{n=m+1}^{\infty} \frac{b_n}{3^n} \geq \frac{1}{3^r} - [\sum_{n=r+1}^{m-1} \frac{2}{3^n} + \frac{1}{3^m}] > 0$. Contradiction.

Case 2: $a_r > b_r$. Then $0 = \sum_{n=1}^{\infty} \frac{a_n - b_n}{3^n} = \frac{a_r - b_r}{3^r} + \sum_{n=r+1}^{m-1} \frac{a_n - b_n}{3^n} + \frac{a_m - b_m}{3^m} - \sum_{n=m+1}^{\infty} \frac{b_n}{3^n} > \frac{1}{3^r} - [\sum_{n=r+1}^{m-1} \frac{2}{3^n} + \frac{1}{3^m} + \sum_{n=m+1}^{\infty} \frac{2}{3^n}] = \frac{1}{3^r} - \frac{1}{3^r} + \frac{1}{3^m} = \frac{1}{3^m} > 0$. Contradiction. □

Now we are ready to classify the Cantor ternary numbers. Let D be the set of open intervals removed in the Cantor process.

By construction $D = \bigcup_{n=2}^{\infty} \{\bigcup_{g \in A_n} D_{n_g}\} \bigcup D_1$ where $D_{n_g} = (\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}, \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n})$ with $g \in A_n = \{$set of all functions from $\{1, 2, ...n-1\}$ to $\{0, 2\}\}$ represents an arbitrary open interval removed at the nth step of the Cantor process and $D_1 = (\frac{1}{3}, \frac{2}{3})$.

Theorem 5.10. *$x \in C$ if and only if there exist $\{a_n\}$ such that $x = \sum_{1}^{\infty} \frac{a_n}{3^n}$ with $a_n = 0$ or $a_n = 2$.*

Theorem 5.10 is equivalent to the following theorem.

Theorem 5.11. *$x \in D$ if and only if x has no ternary expansion with only 0's and 2's.*

Proof. Assume $x \in D$ then $x \in D_{N_g}$ for some N and g.

Hence $x \in \left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}, \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}\right)$ implies $\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} < x < \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}$.

Now $x = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \sum_{i=n+1}^{\infty} \frac{b_i}{3^i}$ where b_i's cannot be all zeroes or all twos.

For if $b_i = 0$ for all $i > n$, then $x = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}$ which is the left endpoint of $\left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}, \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}\right)$. So if $b_{n+1} = 0$, then there exist $k > n+1$ such that $b_k = 1$ or $b_k = 2$.

Also if $b_i = 2$ for all $i > n$, then $x = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \sum_{i=n+1}^{\infty} \frac{2}{3^i} = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \frac{1}{3^n} = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}$ which is the right endpoint of $\left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}, \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}\right)$. So if $b_{n+1} = 2$, then there exist $k > n+1$ such that $b_k = 0$ or $b_k = 1$.

It follows $0 < \sum_{i=n+1}^{\infty} \frac{b_i}{3^i} < \sum_{i=n+1}^{\infty} \frac{2}{3^i} = \frac{1}{3^n}$.

If x has two ternary expansions, by theorem 5.8, $x = \sum_{i=1}^{m-1} \frac{a_i}{3^i} + \frac{k}{3^m}$ where k=1 or 2. Moreover, by theorem 5.9, $x = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \sum_{j=n+1}^{m-1} \frac{a_j}{3^j} + \frac{k}{3^m}$ and any other expansion of x must agree in the first m-1 terms of the series. Hence every expansion of x has a nonremovable one in the nth position, that is x has no ternary expansion with 0's and 2's.

For the other direction, assume $y = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \sum_{i=n+1}^{\infty} \frac{b_i}{3^i}$ where $0 < \sum_{i=n+1}^{\infty} \frac{b_i}{3^i} < \sum_{i=n+1}^{\infty} \frac{2}{3^i} = \frac{1}{3^n}$. It follows $\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} < y < \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}$. Hence $y \in D_{N_g}$. □

As has been shown by the preceding theorem, the Cantor set C consists of all those real numbers in [0,1] that have ternary expansion (a_n) for which a_n is never 1. If x has two ternary expansions, we put x in the Cantor set if one of the expansions has no term equal to 1.

Before continuing with our proof of the existence of a measurable set that is not Borel, we will take a moment to show that the Cantor set is uncountable. Our proof will use the Schroeder-Bernstein Equivalence Theorem which states: If there exists one-to-one mappings f and g going from $A \to B$ and $B \to A$ respectively, then there exists a bijection from $A \to B$.

Theorem 5.12. *The Cantor set is uncountable.*

Proof. Let I, the identity map, and ϕ be defined as follows.
$I : C \to (0,1)$
$\phi : (0,1) \to C$ where $\phi\left(\sum_{n=1}^{\infty} \frac{a_n}{2^n}\right) = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}$.
If a binary number has two binary expansions, choose the one with nonrepeated ones. Clearly, ϕ is well-defined. ϕ is strictly increasing since if $x, y \in (0,1)$ and $x < y$, then $\phi(x) < \phi(y)$. Hence ϕ is one-to-one. Since I is one-to-one and ϕ is one-to-one, then by the Schroeder-Bernstein Equivalence Theorem, C and (0,1) have the same cardinality. Since (0,1) is uncountable, C is uncountable. □

5.3 The Cantor Ternary Function

Definition 5.13. *The Cantor ternary function. Let x be a real number in [0,1] with the ternary expansion (a_n), $x = \sum_1^{\infty} \frac{a_n}{3^n}$ with $a_n = 0$, 1 or 2. Let $N = \infty$ if none of the a_n are 1, and otherwise let N be the smallest value of n such that $a_n = 1$. Let $b_n = \frac{1}{2}a_n$ for $n < N$ and $b_N = 1$. The function f defined by setting $f(x) = \sum_{n=1}^{N} \frac{b_n}{2^n}$ is called the Cantor ternary function.*

The next lemma shows that the Cantor ternary function is well defined.

Lemma 5.14. *If x has two ternary expansions, then $\sum_{n=1}^{N} \frac{b_n}{2^n}$ is independent of the ternary expansion of x.*

Proof. By theorem 5.8 we know that if x has two ternary expansions then $x = \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{k}{3^m}$ where $k = 1$ or $k = 2$ and $a_n \in \{0, 1, 2\}$.

Case 1: If k=1, then let the 1st expansion of x be $x = \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m}$ and the second expansion be $x = \sum_{n=1}^{m-1} \frac{a_n}{3^n} + \sum_{n=1}^{\infty} \frac{2}{3^{m+n}}$.

For the first expansion: If N=m, then $f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m}) = \sum_{n=1}^{m-1} \frac{b_n}{2^n} + \frac{1}{2^m}$; and if $N < m$, then $f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m}) = \sum_{n=1}^{N-1} \frac{b_n}{2^n} + \frac{1}{2^N}$.

Now for the second expansion: If $N = \infty$, then $f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \sum_{n=1}^{\infty} \frac{2}{3^{m+n}}) = \sum_{n=1}^{m-1} \frac{b_n}{2^n} + \sum_{n=1}^{\infty} \frac{1}{2^{m+n}} = \sum_{n=1}^{m-1} \frac{b_n}{2^n} + \frac{1}{2^m} = f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m})$; and if $N < m$, then $f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \sum_{n=1}^{\infty} \frac{2}{3^{m+n}}) = \sum_{n=1}^{N-1} \frac{b_n}{2^n} + \frac{1}{2^N} = f(\sum_{n=1}^{m-1} \frac{a_n}{3^n} + \frac{1}{3^m})$
Hence $\sum_{n=1}^{N} \frac{b_n}{2^n}$ is independent of the ternary expansion of x when k=1.

Case 2: If k=2, then proof follows mutandis mutates as above. □

The aim of this section is to show that the Cantor ternary function is continuous and monotonically increasing, but in order to do this we need the following theorems.

Theorem 5.15. *The Cantor function is constant on each of the open intervals removed in the Cantor process. Moreover, the value of the constant on the interval (a,b) removed is equal to $f(a) = f(b)$.*

Proof. Let $x \in D_{n_g}$ then $x = \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n} + \sum_{i=n+1}^{\infty} \frac{b_i}{3^i}$. It follows $f(x) = \sum_{i=1}^{n-1} \frac{\frac{1}{2}g(i)}{2^i} + \frac{1}{2^n}$. Hence f is constant on D_{n_g} and it equals $f\left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}\right) = f(x) = f\left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}\right)$. □

Theorem 5.16. *The Cantor function is monotonically increasing on the Cantor set.*

Proof. Let $x, y \in C$ such that $x < y$. Let $x = \sum_{i=1}^{m-1} \frac{a_i}{3^i} + \frac{a_m}{3^m} + \sum_{i=m+1}^{\infty} \frac{b_i}{3^i}$ and $y = \sum_{i=1}^{m-1} \frac{c_i}{3^i} + \frac{c_m}{3^m} + \sum_{i=m+1}^{\infty} \frac{d_i}{3^i}$ where $\{a_i\}$, $\{b_i\}$, $\{c_i\}$, $\{d_i\}$ are sequences of 0's and 2's. There exists a smallest m such that $a_i = c_i$ for $i < m$ and $a_m = 0$ and $c_m = 2$.

If $b_i = 2$ for all $i > m$, and $d_i = 0$ for all $i > m$ then $x = \sum_{i=1}^{m-1} \frac{a_i}{3^i} + \frac{0}{3^m} + \sum_{i=m+1}^{\infty} \frac{2}{3^i}$ and $y = \sum_{i=1}^{m-1} \frac{a_i}{3^i} + \frac{2}{3^m}$. It follows $f(x) = \sum_{i=1}^{m-1} \frac{\frac{1}{2}a_i}{2^i} + \frac{0}{2^m} + \sum_{i=m+1}^{\infty} \frac{\frac{1}{2}2}{2^i} = \sum_{i=1}^{m-1} \frac{\frac{1}{2}a_i}{2^i} + \frac{0}{2^m} + \frac{1}{2^m} = \sum_{i=1}^{m-1} \frac{\frac{1}{2}a_i}{2^i} + \frac{1}{2^m} = f(y)$.

Otherwise,
$f(x) = \sum_{i=1}^{m-1} \frac{\frac{1}{2}a_i}{2^i} + \frac{0}{2^m} + \sum_{i=m+1}^{\infty} \frac{\frac{1}{2}b_i}{2^i} < \sum_{i=1}^{m-1} \frac{\frac{1}{2}a_i}{2^i} + \frac{1}{2^m} + \sum_{i=m+1}^{\infty} \frac{\frac{1}{2}d_i}{2^i} = f(y)$
since $\sum_{i=m+1}^{\infty} \frac{\frac{1}{2}b(i)}{2^i} < \sum_{i=m+1}^{\infty} \frac{\frac{1}{2}2}{2^i} = \frac{1}{2^m}$. □

Theorem 5.17. *The Cantor function is monotonically increasing on [0,1].*

Proof. Let $x, y \in [0,1]$ such that $x < y$.

Case 1: Let $x, y \in C$ then $f(x) \leq f(y)$ since the Cantor function is monotonically increasing on the Cantor set.

Case 2: Let $x, y \in D$. Let $x \in (a,b)$ and $y \in (c,d)$ with $(a,b), (c,d) \subset D$. Now $x < y$ implies $a \leq c$. Moreover $f(x) = f(a)$ and $f(y) = f(c)$ implies $f(x) \leq f(y)$.

Case 3: Let $x \in C$ and $y \in D$. Let $y \in (a,b) \subset D$. Now $x < y$ implies $x \leq a$. Since x and a are Cantor numbers, it follows $f(x) \leq f(a)$. Moreover $f(y) = f(a)$ by theorem 5.15. Hence $f(x) \leq f(y)$.

Case 4: Let $x \in D$ and $y \in C$. Let $x \in (a,b) \subset D$. Now $x < y$ implies $b \leq y$. Since b and y are Cantor numbers, it follows $f(b) \leq f(y)$. Moreover $f(x) = f(b)$ by theorem 5.15. Hence $f(x) \leq f(y)$. □

Theorem 5.18. *If $f : [a,b] \to \mathbb{R}$ is monotonically increasing then f is continuous if and only if for all y such that $f(a) \leq y \leq f(b)$, there exist $x \in [a,b]$ such that $f(x) = y$.*

Proof. Assume f is continuous, then by the intermediate value theorem there exist an x such that $f(x) = y$ for all $y \in [f(a), f(b)]$.

Conversely, assume $x \in [a,b]$ such that $f(x) = y$, for all y where $f(a) < y < f(b)$. Let $\epsilon > 0$ and let y_1 and $y_2 \in [f(a), f(b)]$ such that $f(a) < y_1 < y < y_2 < f(b)$ where $y - y_1 < \epsilon$ and $y_2 - y < \epsilon$. Let x_1 and $x_2 \in [a,b]$ such that $x_1 < x < x_2$, with $f(x_1) = y_1$, and $f(x_2) = y_2$. Let $\delta = min\{|x_1 - x|, |x_2 - x|\}$ then for all $c \in (x - \delta, x + \delta)$ we have $|c - x| < \delta$. We want $|f(c) - f(x)| < \epsilon$.

Case 1: $x_1 < c \leq x$ implies $f(x_1) \leq f(c) \leq f(x)$.
Now $f(x) - f(x_1) = f(x) - f(c) + f(c) - f(x_1)$. Since$f(x) - f(x_1) < \epsilon$, it follows $f(x) - f(c) + f(c) - f(x_1) < \epsilon$. Moreover, $f(x) - f(c) \geq 0$ and $f(c) - f(x_1) \geq 0$, implies $f(x) - f(c) < \epsilon$.

Case 2: $x \leq c \leq x_2$ implies $f(x) \leq f(c) \leq f(x_2)$.
Now $f(x_2) - f(x) = f(x_2) - f(c) + f(c) - f(x)$. Also $f(x_2) - f(x) < \epsilon$ implies $f(x_2) - f(c) + f(c) - f(x) < \epsilon$. Moreover, $f(x_2) - f(c) \geq 0$ and $f(c) - f(x) \geq 0$ implies $f(c) - f(x) < \epsilon$.
Case 1 and 2 imply $|f(c) - f(x)| < \epsilon$. □

Corollary 5.19. *The Cantor function is continuous.*

Proof. Let $y \in [f(0), f(1)]$ this implies $0 \leq y \leq 1$. It follows $y = \sum_{i=1}^{\infty} \frac{a_i}{2^i}$ where $a_i = 0$ or $a_i = 1$. Let $x = \sum_{i=1}^{\infty} \frac{2a_i}{3^i}$, then $x \in C$. It follows that for all $y \in [0,1]$, there exist $x \in C$ such that $f(x) = y$. Hence f is continuous by previous theorem. □

5.4 A Measurable Set that is not Borel

Let f_1 be the Cantor function, and define f by $f(x) = f_1(x) + x$.
The proof of the existence of a measurable set that is not Borel requires that f map the

Cantor set onto a set of positive measure. But in order to show that f maps the Cantor set onto a set of positive measure, we need to show that f is a homeomorphism of [0,1] onto [0,2]. The proof follows.

Recall that a space X is compact if every open covering of X contains a finite subcollection that also covers X [Mun00].

Theorem 5.20. *f is a homeomorphism of [0,1] onto [0,2].*

In order to prove this theorem we need the following result. Let $f : X \to Y$ be a bijective continuous function. If X is compact and Y a metric space, then f is a homeomorphism.

Proof. Let $x, y \in [0,1]$ such that $x < y$, then $f(x) = f_1(x) + x$ and $f(y) = f_1(y) + y$. Now $f_1(x) \leq f_1(y)$ since the Cantor function is monotonically increasing. Since $x < y$, it follows $f_1(x) + x < f_1(y) + y$. Hence f is one-to-one, since f is strictly increasing.

Now f_1 and the identity function are continuous. Since the sum of two continuous functions is continuous, f is continuous.

Since $f(0) = f_1(0) + 0 = 0$, $f(1) = f_1(1) + 1 = 2$ and f is continuous, it follows f takes all the values between [0,2]. Hence f is onto.

Since $f : [0,1] \to [0,2]$ is a bijective continuous function with [0,1] compact and [0,2] a metric space, it follows from result above that f is a homeomorphism. □

Theorem 5.21. *f maps the Cantor set onto a set of measure 1.*

Proof. Recall $D = [0,1] - C$ where an arbitrary open interval removed at the nth step of the Cantor process is given by $D_{n_g} = \left(\sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{1}{3^n}, \sum_{i=1}^{n-1} \frac{g(i)}{3^i} + \frac{2}{3^n}\right)$
where $g \in A_n = \{$set of all functions from$\{1, 2, \ldots, n-1\}$ to $\{0, 2\}\}$.
Hence $D = \bigcup_{n=2}^{\infty}\{\bigcup_{g \in A_n} D_{n_g}\} \bigcup D_1$ where $D_1 = (\frac{1}{3}, \frac{2}{3})$.

Now $f(D_{n_g}) = \{f_1(x) + x | x \in D_{n_g}\}$ and $f_1(D_{n_g}) = \{f_1(x) | x \in D_{n_g}\}$. Since f_1 is constant on each D_{n_g}, let $f_1(D_{n_g}) = C_{n_g}$. Then $f(D_{n_g}) = f_1(D_{n_g}) + D_{n_g} = C_{n_g} + D_{n_g}$. Hence $m(f(D_{n_g})) = m(D_{n_g})$ by the translation invariant property on m. Now $f(D) = \bigcup_{n=2}^{\infty}\{\bigcup_{g \in A_n} f(D_{n_g})\} \cup f(D_1)$ is clear by definition. Also since D_{n_g} and D_{m_h} are disjoint sets and f is one-to-one, $f(D_{n_g}) \cap f(D_{m_h}) = \emptyset$ if $g \neq h$ or $n \neq m$.

Therefore,
$m(f(D)) = m\left(\bigcup_{n=2}^{\infty}\{\bigcup_{g \in A_n} f(D_{n_g})\} \cup f(D_1)\right) = \sum_{n=1}^{\infty} \sum_{g \in A_n} m(f(D_{n_g})) + m(f(D_1)) =$

$\sum_{n=1}^{\infty} \sum_{g \in A_n} m(D_{n_g}) + m(D_1) = m(D) = 1$ by the countable additivity property of Lebesgue measure. Now $[0,1] = D \cup C$ and $f([0,1]) = f(D \cup C) = f(D) \cup f(C) = [0,2]$. Therefore, $2 = m([0,2]) = m(f(D) \cup f(C)) = m(f(D)) + m(f(C)) = 1 + m(f(C))$. Hence $m(f(C)) = 1$. □

Now we are ready to prove the existence of a measurable set that is not Borel.

Theorem 5.22. *Let $g = f^{-1}$. There is a measurable set A such that $g^{-1}(A)$ is not measurable. In particular, there exists a measurable set which is not a Borel set.*

Proof. Note $f(D)$ is measurable since $f(D_{n_g})$ and $f(D_1)$ are measurable and $f(D) = \bigcup_{n=2}^{\infty}\{\bigcup_{g \in A_n} f(D_{n_g})\} \bigcup f(D_1)$. Moreover, $f(C)$ is measurable since $f(C) = [0,2] - f(D)$. $f(C)$ measurable and $m(f(C)) = 1$ imply there exists a nonmeasurable set $A \subset f(C)$.

Let $g = f^{-1}$. Now g is continuous and by theorem 5.3 it is measurable. Let $B = f^{-1}(A) \subset C$ then B is measurable since $m(B) \leq m(C) = 0$. However, $g^{-1}(B) = f(B) = A$ is not measurable. It follows B is not a Borel set since the inverse image of a Borel set is measurable for measurable functions. □

The following is an example that shows that the composition of two measurable functions may not be measurable. But first recall that the characteristic function of B is defined as follows: $X_B(x) = 1$ if $x \in B$ and $X_B(x) = 0$ if $x \notin B$.

Let $f(x) = f_1(x) + x$ where f_1 is the Cantor function. By theorem 5.22, there exists a measurable set B such that $f(B)$ is not measurable. Let $g = f^{-1}$ and $h = X_B$, then $(h \circ g)(x) = h(g(x)) = X_B(f^{-1}(x))$. Note that $X_B(f^{-1}(x)) > \frac{1}{2}$ if and only if $f^{-1}(x) \in B$. Now $\{x|(h \circ g)(x) > \frac{1}{2}\} = \{x|X_B(f^{-1}(x)) > \frac{1}{2}\} = \{x|f^{-1}(x) \in B\} = \{x|x \in f(B)\} = f(B)$. It follows $h \circ g$ is not measurable since $f(B)$ is not measurable.

Chapter 6

A Necessary and Sufficient Condition for Riemann Integrability

We are familiar with the definition that a function is Riemann integrable if its upper and lower integral agree. In this chapter we will learn another way of classifying a Riemann integrable function. We will prove that a function f is Riemann integrable precisely when the set of points at which f is discontinuous has measure zero [Roy88]. Sections 6.1 through 6.4 contain the theorems necessary to prove this.

6.1 Properties of Measurable Functions

First we will prove that the supremum and the infimum of a sequence of measurable functions is measurable, but we need the lemma below to do that.

Lemma 6.1. *Let $\{f_n\}$ be a sequence of measurable functions. Let $g = supf_n$ then* $\{x : g(x) > \alpha\} = \bigcup_{n=1}^{\infty}\{x : f_n(x) > \alpha\}$

Proof. Let $E = \{x : g(x) > \alpha\}$ and $E_n = \{x : f_n(x) > \alpha\}$. Let $x_1 \in E$ then $g(x_1) > \alpha$. Since $g(x_1) = supf_n(x_1)$, $supf_n(x_1) > \alpha$. It follows there exist N such that $f_N(x_1) > \alpha$, i.e. $x_1 \in \{x : f_N(x) > \alpha\} = E_N$. Hence, $E \subset \bigcup_{n=1}^{\infty} E_n$.

Let $x_1 \in E_n$ then $f_n(x_1) > \alpha$. Since $g = supf_n$, $g(x_1) \geq f_n(x_1) > \alpha$ for all n. Hence $g(x_1) > \alpha$ implies $x_1 \in E$. Hence $\{x : g(x) > \alpha\} = \bigcup_{n=1}^{\infty}\{x : f_n(x) > \alpha\}$. □

Lemma 6.2. *Let $\{f_n\}$ be a sequence of measurable functions then the $supf_n$ and the $inf f_n$ are measurable.*

Proof. We know from lemma 6.1 that $\{x : g(x) > \alpha\} = \bigcup_{n=1}^{\infty}\{x : f_n(x) > \alpha\}$. Since each f_n is measurable and the countable union of measurable sets is measurable, g is measurable. The proof for $inf f_n$ is done similarly. □

Definition 6.3. *A property is said to hold almost everywhere (abbreviated a.e.) if the set of points where it fails to hold is a set of measure zero.*

Lemma 6.4. *If f is a measurable function and f=g a.e., then g is measurable.*

Proof. $\{x : g(x) > \alpha\} = \{x : f(x) > \alpha\} - \{x : g(x) \leq \alpha$ and $f(x) > \alpha\} \cup \{x : f(x) \leq \alpha$ and $g(x) > \alpha\}$ is measurable since $\{x : f(x) > \alpha\}$ is measurable, $\{x : g(x) \leq \alpha$ and $f(x) > \alpha\}$ and $\{x : f(x) \leq \alpha$ and $g(x) > \alpha\}$ are both measurable with measure zero and measurable sets form a sigma algebra. □

Recall that a step function, φ, is a linear combination of characteristic functions, i.e. $\varphi = \sum_{i=1}^{n} a_i X_{E_i} + \sum_{i=1}^{k} b_i X_{e_i}$, where E_i's are intervals and $e_i \in \mathbb{R}$ [Rud87]. We will show that every step function is measurable, but first we need the lemma below.

Lemma 6.5. *Let $h = aX_I$ and $g = bX_{\{c\}}$, where I is an interval and $c \in \mathbb{R}$, then h and g are measurable.*

Proof. For any $r \in \mathbb{R}$ we have $\{x|h(x) > r\} = I$ if $a > r$, and $\{x|h(x) > r\} = \emptyset$ if $a \leq r$. Moreover $\{x|g(x) > r\} = c$ if $b > r$ and $\{x|g(x) > r\} = \emptyset$ if $b \leq r$. Clearly h and g are measurable. □

Theorem 6.6. *Every step function is measurable.*

Proof. Let f be a step function, then $f = \sum_{i=1}^{n} a_i X_{I_i} + \sum_{i=1}^{k} b_i X_{\{c_i\}}$ where $I_i = (\alpha_i, \beta_i)$, with I_i's disjoint and $c_i \notin I_j$ for all i and j. By lemma 6.5, we know each $a_i X_{I_i}$ and $b_i X_{\{c_i\}}$ is measurable. Now $a_1 X_{I_1} + a_2 X_{I_2}$ is measurable since the sum of two measurable functions is measurable. Assume $a_1 X_{I_1} + a_2 X_{I_2} + \cdots + a_{n-1} X_{I_{n-1}}$

is measurable. Now $(a_1X_{I_1} + a_2X_{I_2} + \cdots + a_{n-1}X_{I_{n-1}}) + a_nX_{I_n}$ is measurable since $(a_1X_{I_1} + a_2X_{I_2} + \cdots + a_{n-1}X_{I_{n-1}})$ and $a_nX_{I_n}$ are measurable and the sum of two measurable functions is measurable.

In a similar way we can show that $b_1X_{\{c_1\}} + b_2X_{\{c_2\}} + \cdots + b_kX_{\{c_k\}}$ is measurable. Moreover f is measurable since f is the sum of two measurable functions. □

6.2 Extending Riemann Integration

The key to extending Riemann integration lies in proposition 6.11 which shows that a bounded measurable function on a closed interval [a,b] and simple functions are a natural extension of Riemann integrable functions on [a,b] and step functions. But in order to prove proposition 6.11 we need the following.

Definition 6.7. *The integral of a characteristic function is defined as follows* $\int X_E = m(E)$.

Definition 6.8. *Simple Function* $\varphi = \sum_{i=1}^{n} a_iX_{E_i}$ *where* E_i*'s are measurable sets.*

Definition 6.9. *The integral of a simple function is defined as follows* $\int \sum_{i=1}^{n} a_iX_{E_i} = \sum_{i=1}^{n} a_im(E_i)$.

Proposition 6.10. *Let* φ *and* ψ *be simple functions*
a) $\int(a\varphi + b\psi) = a\int\varphi + b\int\psi$
b) If $\varphi \geq \psi$ *a.e., then* $\int\varphi \geq \int\psi$

Proposition 6.11. *Let f be defined and bounded on a measurable set E with mE finite. In order that* $inf \int_E \psi(x)dx = sup \int_E \varphi(x)dx$ *for all simple functions* $\psi(x) \geq f(x)$ *and* $\varphi(x) \leq f(x)$*, it is necessary and sufficient that f be measurable.*

Proof. Assume f is measurable and bounded by M. We will first partition the range of f as follows. Let $E_k = \{x : \frac{(k-1)M}{n} < f(x) \leq \frac{kM}{n}; -n \leq k \leq n\}$ and $M = sup_{x\in E}|f(x)|$. Now E_k's are measurable, disjoint and have union E. Thus $\sum_{k=-n}^{n} m(E_k) = m(E)$.

Create simple functions ψ and φ on the sets E_k such that $\varphi \leq f \leq \psi$.
Let $\psi_n = \frac{nM}{n}X_{E_n} + \frac{(n-1)M}{n}X_{E_{n-1}} + \cdots + \frac{kM}{n}X_{E_k} + \cdots + \frac{-nM}{n}X_{E_{-n}}$
and $\varphi_n = \frac{(n-1)M}{n}X_{E_n} + \frac{(n-2)M}{n}X_{E_{n-1}} + \cdots + \frac{(k-1)M}{n}X_{E_{k-1}} + \cdots + \frac{(-n-1)M}{n}X_{E_{(-n-1)}}$.
It follows $\psi_n = \frac{M}{n}\sum_{k=-n}^{n} kX_{E_k}(x)$, and $\varphi_n = \frac{M}{n}\sum_{k=-n}^{n}(k-1)X_{E_k}(x)$.

Now $\int \psi_n - \varphi_n = \int \frac{M}{n}(X_{E_n} + X_{E_{n-1}} + \cdots + X_{E_{-n+1}} + X_{E_{-n}})$
$= \frac{M}{n} \int \sum_{k=-n}^{n} X_{E_k} = \frac{M}{n} \int X_{\bigcup_{k=-n}^{n} E_k} = \frac{M}{n} \int X_E = \frac{M}{n} m(E)$.
Moreover $sup \int \psi \geq \int \varphi_n$ and $inf \int \psi \leq \int \psi_n$.
Also $\int \psi_n > \int \varphi_n$ since $\varphi \leq f \leq \psi$. Hence $\int \psi_n \geq inf \int \psi \geq sup \int \varphi \geq \int \varphi_n$.
Since $sup \int \varphi \geq \int \varphi_n$, then $\frac{M}{n} m(E) = \int \psi_n - \int \varphi_n \geq \int \psi_n - sup \int \varphi \geq inf \int \psi - sup \int \varphi \geq 0$ for all n. Hence $inf \int \psi = sup \int \varphi$.

Now for the other direction, let f be some fixed bounded function. Suppose $\alpha = sup_{\varphi \leq f} \int_E \varphi(x) = inf_{f \leq \psi} \int_E \psi(x)$ where φ and ψ are simple functions. Hence for all n there exist $\varphi_n \leq f$ and $\psi_n \geq f$ such that $\int_E \varphi_n > \alpha - \frac{1}{n}$ and $\int_E \psi_n < \alpha + \frac{1}{n}$. It follows $\int_E \psi_n - \int_E \varphi_n = \int_E \psi_n - \varphi_n < (\alpha + \frac{1}{n}) - (\alpha - \frac{1}{n}) = \frac{2}{n}$.

Let $\psi^* = inf\ \psi_n$ and $\varphi^* = sup\ \varphi_n$. Now ψ^* and φ^* are measurable and bounded by lemma 6.2. Let $g = \psi^* - \varphi^*$. We have that $g \geq 0$ and it is bounded and measurable. Moreover $g = \psi^* - \varphi^* \leq \psi_n - \varphi_n$.

Suppose $m(\{x : \psi^*(x) - \varphi^*(x) > 0\}) > 0$ then there exist $k > 0$ such that $m(\{x : \psi^*(x) - \varphi^*(x) > \frac{1}{k}\}) > 0$. Let $E_k = \{x : \psi^*(x) - \varphi^*(x) > \frac{1}{k}\}$ and $\varphi = \frac{1}{k} X_{E_k}$. It follows $\varphi \leq g$ since $g = \psi^* - \varphi^* > \frac{1}{k}$ for $x \in E_k$. Moreover $\int \varphi = \frac{m(E_k)}{k} > 0$. Hence $sup_{\varphi \leq g} \int_E \varphi(x) > 0$. Since g is measurable and bounded we have by part a that $sup_{\varphi \leq g} \int_E \varphi(x) = inf_{g \leq \psi} \int_E \psi(x)$.

Let $\psi = \psi_n - \varphi_n$. It follows $\psi \geq g$ and $\int \psi = \int (\psi_n - \varphi_n) \leq \frac{2}{n}$. Hence $inf_{g \leq \psi} \int_E \psi(x) \leq \frac{2}{n}$ for all n implies $inf_{g \leq \psi} \int_E \psi(x) = 0$. So we have $sup \int \varphi > 0$ and $inf \int \psi = 0$ which contradicts $sup \int \varphi = inf \int \psi$. Hence $m(\{x : \psi^*(x) - \varphi^*(x) > 0\}) = 0$. It follows $\psi^* = \varphi^*$ a.e.. Now $\varphi^* \leq f \leq \psi^*$ implies $\varphi^* = f$ a.e.. Hence f is measurable. □

As was shown in the proposition above, Lebesgue's construction of the integral is different from Riemann's construction in that it partitions the range of the function rather than the domain[Bur98].

Definition 6.12. *If f is a bounded measurable function defined on a measurable set E with mE finite, we define the (Lebesgue) integral of f over E by $\int_E f(x)dx = inf \int_E \psi(x)dx$ for all simple functions $\psi(x) \geq f(x)$.*

Having defined the Lebesgue integral of a function f, we will show in the next proposition that if f is Riemann integrable then the Riemann integral of f equals the Lebesgue integral of f. Recall that the upper Riemann integral of f is defined by $R\int_a^{\bar{b}} f(x)dx = inf \int_a^b \psi(x)dx$ for all step functions $\psi(x) \geq f(x)$. And the lower Riemann integral of f is defined by $R\int_{\underline{a}}^b f(x)dx = sup \int_a^b \varphi(x)dx$ for all step functions $\varphi(x) \leq f(x)$. Hence a function f is Reimann integrable on [a,b] if the upper integral is equal to the lower integral. The Reimann integral of f is denoted by $R\int_a^b f(x)dx$ [Bar66].

Proposition 6.13. *Let f be a bounded function defined on [a,b]. If f is Riemann integrable on [a,b], then it is measurable and $R\int_a^b f(x)dx = \int_a^b f(x)dx$.*

Proof. Let f be a Riemann integrable function.

Now we have that $inf \int_{\psi \geq f} \psi^{step} \geq inf \int_{\psi \geq f} \psi^{simple}$. Moreover $sup \int_{\varphi \leq f} \varphi^{step} \leq sup \int_{\varphi \leq f} \varphi^{simple}$, and $sup \int_{\varphi \leq f} \varphi^{simple} \leq inf \int_{\psi \geq f} \psi^{simple}$.
Hence $sup \int \varphi^{step} \leq sup \int \varphi^{simple} \leq inf \int \psi^{simple} \leq inf \int \psi^{step}$.

Since f is Riemann integrable, $sup \int_{\varphi \leq f} \varphi^{step} = inf \int_{\psi \geq f} \psi^{step}$ and it follows $sup \int \varphi^{simple} = inf \int \psi^{simple}$. Hence f is measurable. □

Proposition 6.13 shows that every Riemann integrable function is measurable. However, the converse is not true. Consider the measurable set $Q \cap [0,1]$. It follows that $f = X_{Q\cap[0,1]}$ is measurable. However, f is not Riemann integrable since we have $R\int_a^{\bar{b}} f = 1$ and $R\int_{\underline{a}}^b f = 0$.

6.3 Properties of Lebesgue Integration

Theorem 6.14. *Given $f \geq 0$, bounded and measurable, Lebesgue integral of f is zero if and only if f equals zero a.e..*

Proof. Assume $E_0 = \{x : f(x) > 0\}$ then $E_0 = \bigcup_{n=1}^{\infty}\{x : f(x) > \frac{1}{n}\}$.
Let $E_n = \{x : f(x) > \frac{1}{n}\}$. Suppose that $m(E_0) > 0$, then there exists N such that $m(E_N) > 0$. Let $\varphi_N = \frac{1}{N}X_{E_N}$ by design $\varphi \leq f$. Now $\int \varphi_N = \frac{1}{N}\int X_{E_N} = \frac{1}{N}(m(E_N)) > 0$. Since $\int f = sup \int_{\varphi \leq f} \varphi(x)$, where φ is a simple function, it follows $0 < \int \varphi_N(x) \leq sup \int_{\varphi \leq f} \varphi(x) = \int f$. Contradiction. Hence $m(E_0) = m(\{x : f(x) > 0\}) = 0$.

Conversely, let φ be a simple function such that $\varphi \leq f$ and $\varphi > 0$. Let $\varphi = \sum_{n=1}^{\infty} a_n X_{E_n}$ where $a_n > 0$ and $E_n \cap E_m = \emptyset$ for all n, m where $n \neq m$. Now $\int \varphi(x) = \int \sum_{n=1}^{\infty} a_n X_{E_n} = \sum_{n=1}^{\infty} a_n m(E_n)$. Since $\varphi \leq f$, then $a_n = \varphi(x) \leq f(x)$ for

$x \in E_n$. Hence $E_n \subset \{x : f(x) \geq a_n\}$. Moreover, $a_n > 0$ implies $E_n \subset \{x : f(x) \geq a_n\} \subset \{x : f(x) > 0\}$. Hence $m(E_n) \subset m(\{x : f(x) \geq a_n\}) \subset m(\{x : f(x) > 0\})$ by proposition 2.4. Since $m(\{x : f(x) > 0\}) = 0$, then $m(E_n) = 0$ for all n. It follows $\int \varphi(x) = 0$. Hence $\int f = sup \int_{\varphi \leq f} \varphi(x) = \int \varphi(x) = 0$ since φ is an arbitrary simple function below f. □

The following are some properties about bounded measurable functions defined on a set of finite measure that will be used in later proofs.

Proposition 6.15. *If f and g are bounded measurable functions defined on a set E of finite measure, then:*
i. $\int_E (af + bg) = a \int_E f + b \int_E g$.
ii. If f=g a.e., then $\int_E f = \int_E g$.
iii. If $f \leq g$ *a.e., then* $\int_E f \leq \int_E g$.
iv. If A and B are disjoint measurable sets of finite measure, then $\int_{AUB} f = \int_A f + \int_B f$.

Next we will prove the Bounded Convergence Theorem, its proof uses the following proposition known as one of Littlewood's theorems [Rud87].

Proposition 6.16. *Let E be a measurable set of finite measure, and $\{f_n\}$ a sequence of measurable functions defined on E. Let f be a real-valued function such that for each x in E we have $f_n \to f(x)$. Then given $\epsilon > 0$ and $\delta > 0$, there is a measurable set $A \subset E$ with $mA < \delta$ and an integer N such that for all $x \notin A$ and all $n \geq N$, $|f_n(x) - f(x)| < \epsilon$.*

Proposition 6.17. *Bounded Convergence Theorem (BCT): Let $\{f_n\}$ be a sequence of measurable functions defined on a set E of finite measure, and suppose that there is a real number M such that $|f_n(x)| \leq M$ for all n and all x. If $f(x) = lim f_n(x)$ for each x in E, then* $\int_E f = \int_E lim f_n = lim \int_E f_n$.

Proof. Let $\epsilon = \delta > 0$ then there exists N and A such that $m(A) < \epsilon$ and $|f_n(x) - f(x)| < \epsilon$ for $n \geq N$ and $x \in E - A$ by proposition 6.16. Now $|\int_E (f - f_n)| \leq \int_E |f - f_n| = \int_{E-A} |f - f_n| + \int_A |f - f_n|$. By design $|f(x) - f_n(x)| < \epsilon$ for $n \geq N$ and $x \in E - A$. Hence $\int_{E-A} |f - f_n| \leq \int_{E-A} \epsilon = m(E - A)\epsilon \leq m(E)\epsilon$ and $\int_A |f - f_n| \leq (2M)m(A) \leq (2M)\epsilon$ since $|f - f_n| \leq |f| + |f_n| \leq 2M$. Therefore, when $n \geq N$ we have $|\int_E f - f_n| \leq \epsilon[m(E) + 2M]$, i.e. $lim |\int_E f - f_n| = 0$ that is $lim \int_E f_n = \int f$. □

6.4 A Necessary and Sufficient Condition for Riemann Integrability

Definition 6.18. *Upper and lower envelopes of a function. Let f be a real-valued function defined on [a,b]. We define the lower envelope h of f to be the function h defined by* $h(y) = sup_{\delta>0} inf_{|x-y|<\delta} f(x)$, *and the upper envelope H by* $H(y) = inf_{\delta>0} sup_{|x-y|<\delta} f(x)$.

In order to prove theorem 6.24 which states that a bounded function f on [a,b] is Riemann integrable if and only if the set of points at which f is discontinuous has measure zero, it is sufficient to show the following. First we need to show that the upper Riemann integral of f equals the Lebesgue integral of the upper envelope of f (theorem 6.21); as well as the lower Riemann integral of f equals the Lebesgue integral of the lower envelope of f (corollary 6.22). Second we need to show that $H(x) = h(x)$ if and only if f is continuous at x (theorem 6.23). We start by proving theorem 6.21, but we need lemma 6.19 and 6.20 to prove it.

Lemma 6.19. *Let f be a bounded function. If* φ *is a step function such that* $\varphi \geq f$ *then* $\varphi \geq H$ *except at a finite number of points.*

Proof. $H(x) = inf_{\delta>0} sup_{|x-y|<\delta} f(y)$ is the upper envelope.

Let $\varphi = \sum_{i=1}^{n} c_i X_{(a_i,b_i)} + d_i X_{\{a_i\}} + e_i X_{\{b_i\}}$. Suppose $x \in (a_i, b_i)$ then there exists δ_x such that $(x - \delta_x, x + \delta_x) \subset (a_i, b_i)$. Now $H(x) \leq sup_{|x-y|<\delta} f(y)$ for any $\delta < 0$. Also $f(y) \leq c_i$ for all y such that $|x - y| < \delta_x$.

It follows $sup_{|x-y|<\delta_x} f(y) \leq sup_{|x-y|<\delta_x} c_i$. Therefore $sup_{|x-y|<\delta_x} f(y) \leq c_i$. Since $H(x) \leq sup_{|x-y|<\delta} f(y)$, it follows $H(x) \leq c_i$. Hence $\varphi \geq H$ except for a finite number of points. □

Note that $\varphi \downarrow H$ means a decreasing sequence converging to H.

Lemma 6.20. *Let f be a bounded function. There exists a sequence* $\{\varphi_n\}$ *of step functions such that* $\varphi \downarrow H$ *where* $\{\varphi_n\}$ *is uniformly bounded.*

Proof. Consider $[a, b] \subset \bigcup_{x\in[a,b]} N(x, \frac{1}{n})$. By compactness of $[a, b]$ there exists $\{x_{1n}, \ldots x_{kn}\}$ such that $[a, b] \subset \bigcup_{i=1}^{k} N(x_{in}, \frac{1}{n})$. Without loss of generality we may assume $\{x_{1n}, \ldots x_{kn}\} \subset \{x_{1m}, \ldots x_{lm}\}$ when $m > n$. Define $\varphi = \sum_{i=1}^{k} sup_{y\in N(x_{in},\frac{1}{n})} f(y) X_{N(x_{in},\frac{1}{n})}$. Now, $\varphi_n \geq H$.

We want to show that $lim\ \varphi_n(x) = H(x)$ for $x \in [a, b]$.

We have that for any $x \in [a, b]$ and n there exist x_{in} such that $x \in N(x_{in}, \frac{1}{n})$. Since $x \in N(x_{in}, \frac{1}{n})$, there exist δx_n such that $N(x, \delta x_n) \subset N(x_{in}, \frac{1}{n})$.

Moreover $sup_{y \in N(x,\delta x_n)} f(y) \leq sup_{y \in N(x_{in},\frac{1}{n})} f(y)$. Note that $\delta x_n \leq \frac{1}{n}$, i.e. lim $\delta x_n = 0$. In particular, $H(x) = lim_{n\to\infty} sup_{y \in N(x,\delta x_n)} f(y)$.

Now there exists k and x_{jk} such that $\frac{1}{k} < \delta x_n$ and $x \in N(x_{jk}, \frac{1}{k}) \subset N(x, \delta x_n)$. Hence $sup_{y \in N(x,\delta x_n)} f(y) \geq sup_{y \in N(x_{jk},\frac{1}{k})} f(y) = \varphi_k(x)$ and $sup_{y \in N(x,\delta x_n)} f(y) \leq sup_{y \in N(x_{in},\frac{1}{n})} f(y) = \varphi_n(x)$. Therefore, $\varphi_k(x) \leq sup_{y \in N(x,\delta x_n)} f(y)$ $\leq \varphi_n(x)$. By design, $\varphi_n(x) \geq \varphi_{n+1}(x)$, that is φ_n is a decreasing sequence. Consequently, $H(x) = lim\ sup_{y \in N(x,\delta x_n)} f(y) = lim\ \varphi_n(x)$. And we have that $|f(x)| \leq M$ implies $|\varphi_n(x)| \leq M$, i.e. φ_n is uniformly bounded. □

Theorem 6.21. *Let f be a bounded function on [a,b] and let H be the upper envelope of f then* $\int_a^b H = R\int_a^{\bar{b}} f$.

Proof. By lemma 6.19 we have $\varphi \geq H$ except for a finite number of points. Hence $\int_a^b H \leq \int_a^b \varphi$ which implies $inf \int_{\varphi \geq f} H \leq inf \int_{\varphi \geq f} \varphi$ which in turn implies $\int H \leq$ $inf \int_{\varphi \geq f} \varphi$. Now $\int_a^b H \leq \int_a^{\bar{b}} f$ since $\int_a^{\bar{b}} f = inf \int_{\varphi \geq f} \varphi$.

By lemma 6.20 we have $|\varphi_n(x)| \leq M$ and $H(x) = lim\ \varphi_n(x)$, then by BCT $\int H(x) = \int lim\ \varphi_n(x) = lim \int \varphi_n(x)$. Since φ_n are decreasing, then we have $lim \int \varphi_n(x) =$ $inf \int \varphi_n(x)$. Hence $inf \int \varphi_n(x) \geq inf \int_{\varphi \geq f} \varphi = \int_a^{\bar{b}} f$. Hence $\int H(x) = lim \int \varphi_n(x) \geq$ $\int_a^{\bar{b}} f$.

Therefore $\int_a^b H \leq \int_a^{\bar{b}} f$ and $\int H(x) \geq \int_a^{\bar{b}} f$ which implies $\int_a^b H = R\int_a^{\bar{b}} f$. □

Corollary 6.22. *Let f be a bounded function on [a,b] and let h be the lower envelope of f then* $\int_a^b h = R\int_{\underline{a}}^b f$.

Proof. Follows mutandis mutates from theorem above.

Theorem 6.23. *Let f be a bounded function on [a,b] with H and h the upper and lower envelope respectively. Then* $H(x) = h(x)$ *if and only if f is continuous at* $x \in [a, b]$.

Proof. Suppose $H(x) = h(x)$ then $inf_{\delta>0} sup_{|x-y|<\delta} f(y) = sup_{\delta>0} inf_{|x-y|<\delta} f(y)$. It follows for any $\frac{\epsilon}{2} > 0$ there exist $\delta > 0$ such that $sup_{|x-y|<\delta} f(y) < H(x) + \frac{\epsilon}{2}$ and

$inf_{|x-y|<\delta} f(y) > H(x) - \frac{\epsilon}{2}$. It follows $sup_{|x-y|<\delta} f(y) - inf_{|x-y|<\delta} f(y) < \epsilon$. Now $sup_{|x-y|<\delta} f(y) - inf_{|x-y|<\delta} f(y) > |f(y) - f(x)|$. Hence $|f(y) - f(x)| < \epsilon$ on $|x-y| < \delta$.

Conversely, suppose f is continuous at $x \in [a, b]$. For all $\epsilon > 0$ there exist $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$. It follows $f(x) - \epsilon < f(y) < f(x) + \epsilon$. Hence $f(x) \leq H(x) = inf_{\delta>0} sup_{|x-y|<\delta} f(y) < f(x) + \epsilon$ and $f(x) \geq h(x) = sup_{\delta>0} inf_{|x-y|<\delta} f(y) \geq f(x) - \epsilon$ implies $f(x) - \epsilon \leq h(x) \leq f(x) \leq H(x) \leq f(x) + \epsilon$, for all $\epsilon > 0$. Hence $H(x) = f(x) = h(x)$. □

Theorem 6.24. *A bounded function f on [a,b] is Riemann integrable if and only if the set of points at which f is discontinuous has measure zero.*

Proof. Assume f is Riemann integrable, then $R\int_a^{\bar{b}} f = R\int_{\underline{a}}^b f$; hence, $\int_a^b H = \int_a^b h$. Now $R\int_a^{\bar{b}} f - R\int_{\underline{a}}^b f = \int_a^b (H - h) = 0$. Since $H - h \geq 0$, by theorem 6.14, $H - h$ is zero a.e. Hence $H(x) = h(x)$ a.e.. Then by theorem 6.23, f is continuous a.e.. This implies that the set of points where f is discontinuous has measure zero.

Conversely, assume f is continuous a.e. then $H(x) = h(x)$ a.e.. It follows $\int H(x) = \int h(x)$. Hence $R\int_a^{\bar{b}} f = R\int_{\underline{a}}^b f$ that is f is Riemann integrable. □

It can sometimes be extremely difficult to determine if a function is Riemann integrable using the usual definition of Riemann integration. For situations like this, theorem 6.24 provides us with a powerful and easier method to determine if a function is Riemann integrable. Let's look at an example.

First, note that the characteristic function of an open set is discontinuous at precisely the boundary of that open set. So if the boundary of the set has measure zero, then the characteristic function is Riemann integrable.

Consider $D(\frac{1}{2})$ which is the open set of the generalized cantor set when $\alpha = \frac{1}{2}$. It follows $X_{D(\frac{1}{2})}$ is discontinuous precisely at $Bd(D(\frac{1}{2}))$. Now $Bd(D(\frac{1}{2})) = C(\frac{1}{2})$. Since $m(C(\frac{1}{2})) = \frac{1}{2}$, $X_{D(\frac{1}{2})}$ is not Riemann integrable on [0,1].

Chapter 7

Convergence Theorems and Applications

One of the great advantages of the Lebesgue integral over the Riemann integral lies in the facilitation of limit operations. Recall that in Riemann integration a sequence of functions $\{f_n\}$ need to converge uniformly for the following equation to be true $lim \int f_k = \int lim f_k$. That is, if $lim f_k = f$ uniformly on [a,b], then $lim \int_a^b f_k(x)dx = \int_a^b f(x)dx = \int_a^b lim f_k(x)dx$ [Rud87]. We will see in this chapter that the Lebesgue integral is more powerful and has greater applications than the Riemann integral.

In chapter 6, we were introduced to the Lebesgue integral for bounded functions on sets of finite measure. We also proved the Bounded Convergence Theorem (BCT) which says that if we have a sequence of bounded measurable functions $\{f_n\}$, on a set E of finite measure and if $lim f_n = f$, then $lim \int_E f_n(x)dx = \int_E f(x)dx = \int_E lim f_n(x)dx$.

We first extend the definition of integral from sets of finite measure to sets of arbitrary measure. We do that in two stages. The first one extends the definition of integral for nonnegative functions.

Definition 7.1. *The integral of a nonnegative function: If f is a nonnegative measurable function defined on a measurable set E, we define $\int_E f = sup_{h \leq f} \int_E h$, where h is a bounded measurable function such that $E_h = \{x : h(x) \neq 0\}$ and $m(E_h) < \infty$.*

Definition 7.2. *A nonnegative measurable function f defined on a measurable set E is integrable if $\int_E f < \infty$.*

Theorem 7.3. *Fatou's Lemma: If $\{f_n\}$ is a sequence of nonnegative measurable functions and $f_n(x) \to f(x)$ a.e. on a set E, then $\int_E f \leq \underline{lim} \int_E f_n$.*

Proof. We will use BCT to prove Fatou's Lemma. Let $h \leq f$ such that h is bounded and $\int_E h = \int_{E_h} h$ with $E_h \subset E$ and $m(E_h) < \infty$. Let $h_n(x) = min\{h(x), f_n(x)\} \leq h(x)$. Hence h_n is measurable, bounded and vanishes outside E_h. In order to use BCT we need $\lim h_n(x)$ to exist for all $x \in E_h$.

Claim $h(x) = lim\ h_n(x)$.

Proof of claim. Case one: $h(x) = f(x)$. We have $lim f_n(x) = f(x)$ and $h_n(x) = min\{f_n(x), h(x)\} = min\{f_n(x), f(x)\}$. Given $\epsilon > 0$ there is N such that $|f_n(x) - f(x)| < \epsilon$ for $n \geq N$. Hence $f(x) - \epsilon \leq f_n(x) \leq f(x) + \epsilon$ implies $f(x) - \epsilon \leq h_n(x) \leq f(x) + \epsilon$ for $n \geq N$. Hence $lim\ h_n(x) = f(x) = h(x)$.

Case two: $h(x) < f(x)$. Let $2\epsilon = f(x) - h(x)$. There is N such that $f(x) - \epsilon < f_n(x) < f(x) + \epsilon$ for $n \geq N$, since $lim f_n(x) = f(x)$. In particular, $h(x) < f(x) - \epsilon < f_n(x)$ for $n \geq N$. Hence $h_n(x) = h(x)$ for $n \geq N$.

Apply BCT, we get that $\int_E h = \int_{E_h} h = lim \int_{E_h} h_n = lim \int_E h_n$. Moreover, $h_n(x) \leq f_n(x)$ implies $\int_E h_n \leq \int_E f_n$, which in turn implies $\underline{lim} \int_E h_n \leq \underline{lim} \int_E f_n$. By design $\underline{lim} \int_E h_n = lim \int_E h_n$. Hence, $\int_E h = lim \int_E h_n \leq \underline{lim} \int_E f_n$. That is $\int_E h \leq \underline{lim} \int_E f_n$. In particular, $\int_E f = sup \int_E h \leq \underline{lim} \int_E f_n$ where $h \leq f$ and h is bounded and vanishes outside a set of finite measure. □

Theorem 7.4. *Monotone Convergence Theorem(MCT): Let $\{f_n\}$ be an increasing sequence of nonnegative measurable functions, and let $f = lim f_n$ a.e.. Then $\int f = lim \int f_n$.*

Note that by $h_n \uparrow h$ we mean an increasing sequence converging to h.

Theorem 7.5. *Fatou's lemma is equivalent to the Monotone Convergence Theorem.*

Proof. Assume $0 \leq f_n$ and $lim f_n(x) = f(x)$ imply $\int f \leq \underline{lim} \int f_n$. Let $h_n \uparrow h$, where $h_n \geq 0$. By Fatou's lemma, $0 \leq \int (h - h_n) \leq \underline{lim} \int (h - h_n) = \underline{lim} \int h + \underline{lim} \int (-h_n) = \int h - \overline{lim} \int h_n$ since $\underline{lim} \int h = \int h$ and $\underline{lim} \int (-h_n) = -\overline{lim} \int h_n$. This implies $\int h \geq \overline{lim} \int h_n$. Moreover by Fatou's $\int h \leq \underline{lim} \int h_n$. Hence we must have $\int h = lim \int h_n$.

Conversely, suppose $0 \leq f_n \uparrow f$ then $lim \int f_n = \int f$. Let $0 \leq h_n$ such that $lim h_n(x) = h(x)$. Let $f_n(x) = inf_{k \geq n} h_k(x) \leq h_n(x)$. Since $f_n \uparrow h$ then by MCT $\int h =$

$\lim \int f_n$. In particular, $f_n \leq h_n$ implies $\lim \int f_n = \underline{\lim} \int f_n \leq \underline{\lim} \int h_n$. Hence, $\int h \leq \underline{\lim} \int h_n$. □

Next, we carry out the second stage of the extension of the definition of integral from sets of finite measure to sets of arbitrary measure. We do that by defining the integral over a set of arbitrary measure for a function that is not necessarily nonnegative.

Definition 7.6. *By the positive part f^+ of a function f we mean the function $f^+ = f \vee 0$; that is, $f^+(x) = max\{f(x), 0\}$. Similarly, we define the negative part f^- by $f^- = (-f) \vee 0$. If f is measurable, so are f^+ and f^-. We have $f = f^+ - f^-$ and $|f| = f^+ + f^-$.*

Definition 7.7. *A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E. In this case we define $\int_E f = \int_E f^+ - \int_E f^-$.*

Theorem 7.8. *Let f be a function defined on a measurable set E. Then f is integrable on E if and only if $|f|$ is integrable.*

Proof. If f is integrable on E then both f^+ and f^- are integrable on E by definition. Conversely if $\int_E |f|$ is finite, so are $\int_E f^+$ and $\int_E f^-$.□

Theorem 7.9. *Lebesgue Dominated Convergence Theorem: Let g be integrable over E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n| \leq g$ on E for almost all x in E we have $f(x) = lim f_n(x)$. Then $\int_E f = \lim \int_E f_n$.*

Proof. Our objective is to show $\underline{\lim} \int f_n = \int f = \overline{\lim} \int f_n$.

By hypothesis, $|f_n| \leq g$ and $\int_E g < \infty$. Let $h_n = g - f_n$ and $h = g - f$. Note $0 \leq g - |f_n|$. Hence $h_n \geq 0$. Moreover, $lim h_n(x) = lim(g(x) - f_n(x)) = g(x) - lim f_n(x) = g(x) - f(x)$.

By Fatou's lemma we have $\int g - \int f = \int g - f = \int h \leq \underline{\lim} \int h_n$ and $\underline{\lim} \int h_n = \underline{\lim} \int g - f_n = \underline{\lim}[\int g - \int f_n] = \underline{\lim} \int g + \underline{\lim}(-\int f_n) = \int g - \overline{\lim} \int f_n$. It follows $\int g - \int f \leq \int g - \overline{\lim} \int f_n$. Hence $\overline{\lim} \int f_n \leq \int f$.

Now let $h_n^* = g + f_n$ and $h^* = g + f$. Note $0 \leq g + |f_n|$. Hence $h_n^* \geq 0$. Moreover, $lim h_n^*(x) = lim(g(x) + f_n(x)) = g(x) + lim f_n(x) = g(x) + f(x)$.

By Fatou's lemma we have $\int g + \int f = \int g + f = \int h^* \leq \underline{\lim} \int h_n^*$ and $\underline{\lim} \int h_n^* = \underline{\lim} \int g + f_n = \underline{\lim}[\int g + \int f_n] = \underline{\lim} \int g + \underline{\lim} \int f_n = \int g + \underline{\lim} \int f_n$. It follows $\int g + \int f \leq \int g + \underline{\lim} \int f_n$. Hence $\int f \leq \underline{\lim} \int f_n$.

Clearly $\underline{lim} \int f_n \leq \overline{lim} \int f_n$. Hence $\overline{lim} \int f_n = \int f = \underline{lim} \int f_n$ which implies $lim \int f_n$ exists and $\int f = lim \int f_n$. □

7.1 Applications

Example 7.10. *An example of a function having an improper Riemann integral without possessing a Lebesgue integral.*

Let $f = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} X_{(n-1,n)}$. Then we have $R\int_1^{\infty} f = lim_{N\to\infty} R\int_1^N f = lim_{N\to\infty} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ which converges . Hence the improper Riemann integral exists.

Now f is integrable if and only if $|f|$ is integrable by theorem 7.7. Moreover, $\int_1^{\infty} |f| = \int_1^{\infty} \sum_{n=1}^{\infty} |\frac{(-1)^n}{n}| X_{(n-1,n)} = \int_1^{\infty} \sum_{n=1}^{\infty} \frac{1}{n} X_{(n-1,n)}$. Let $f_k = \sum_{n=1}^{k} \frac{1}{n} X_{(n-1,n)}$ then $lim f_k(x) = |f(x)|$ and $f_{k+1}(x) \geq f_k(x)$. Hence by MCT $\int_1^{\infty} |f| = lim \int f_k = lim \sum_{n=1}^{k} \frac{1}{n} = \sum_{n=1}^{\infty} \frac{1}{n}$ which diverges. Therefore, f is not Lebesgue integrable.

Example 7.11. *If the Lebesgue and improper Riemann integral of a function exist, then they are equal.*

Suppose f possesses an improper Riemann integral such that f is integrable on the domain.

Case 1: $lim_{x\to c^+} R\int_x^b f = R\int_c^b f$.

Let $\{x_n\} \downarrow c$. Define $f_n = (f)X_{(x_n,b)}$. It follows $|f_n| \leq |f|$ and $lim f_n(x) = f(x)$ on $c < x < b$. Hence by Lebesgue Dominated Convergence Theorem $R\int_c^b f = lim_{n\to\infty} \int_{x_n}^b f = lim_{n\to\infty} \int_c^b f_n = \int_c^b lim f_n = \int_c^b f$. Hence the improper integral equals the Lebesgue integral. All other cases follow mutandis mutates.

Example 7.12. *An example of a sequence of nonnegative Riemann integrable functions that increase monotonically to a bounded function that is not Riemann integrable.*

Let $\{r_i\}$ be the enumeration of rational numbers in [0,1]. Let $\varphi_n = X_{\{\bigcup_{i=1}^n r_i\}}$. By design $\varphi_n \leq \varphi_{n+1}$. Moreover $lim\ \varphi_n = X_{Q\cap[0,1]}$. Now $R\int_0^1 \varphi_n = 0$ for all n.

However, $\overline{\int_0^1} X_{Q\cap[0,1]} = 1$ and $\underline{\int_0^1} X_{Q\cap[0,1]} = 0$. Hence, $X_{Q\cap[0,1]}$ is not Riemann integrable on [0,1]. This example clearly shows that the convergence theorems are theorems for the Lebesgue integral.

Example 7.13. *Consider a sequence of functions whose improper Riemann integrals converge. It is much easier to obtain the convergence with the Lebesgue Dominated Convergence Theorem, than to do it with the usual definition of improper Riemann integral.*

Let us find the $lim_{n\to\infty}\int_0^\infty \frac{sin(e^x)}{1+nx^2}dx$. Let $f_n(x) = \frac{sin(e^x)}{1+nx^2}$ then $|f_n(x)| = \frac{|sin(e^x)|}{1+nx^2} \leq \frac{1}{1+nx^2} \leq \frac{1}{1+x^2} = g(x)$.

Now let $g(x) = \frac{1}{1+x^2}$ then $\int_0^\infty \frac{1}{1+x^2}dx = \frac{\pi}{2}$, that is $g(x)$ is integrable on $[0,\infty)$. Note f_n is dominated by g on $[0,\infty)$. Moreover, $lim_{n\to\infty}f_n(x) = 0$ for all $x \in [0,\infty)$. Hence by Lebesgue Dominated Convergence Theorem $lim_{n\to\infty}\int_0^\infty f_n = \int_0^\infty lim_{n\to\infty}f_n = \int_0^\infty 0 = 0$.

Chapter 8

General Measure Theory

8.1 From Outer Measure to a Measure

In this section, we give the general procedure for obtaining a countably additive measure from an outer measure.

Definition 8.1. *An outer measure μ^* is a nonnegative extended real-valued set function defined on all subsets of a space X with the following properties:*
i. $\mu^*\emptyset = 0$.
ii. $A \subset B \to \mu^*A \leq \mu^*B$, *called monotonicity*
iii. $E \subset \bigcup_{i=1}^{\infty} E_i \to \mu^*E \leq \sum_{i=1}^{\infty} \mu^*E_i$, *called countable subadditivity*

Definition 8.2. *A set E is said to be measurable with respect to μ^* if for every set A we have $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$.*

We will need the following lemma to prove theorem 8.4.

Lemma 8.3. *Let A be any set and $\{E_n\}$ be a finite sequence of disjoint measurable sets. Then $\mu^*(A \cap [\bigcup_{i=1}^{n} E_i]) = \sum_{i=1}^{n} \mu^*(A \cap E_i)$.*

Proof. Clearly true for n=1. Assume it is true for n-1 sets E_i. Now the E_i's are disjoint sets so $(A \cap (\bigcup_{i=1}^{n} E_i)) \cap E_n = A \cap E_n$ and $(A \cap (\bigcup_{i=1}^{n} E_i)) \cap (E_n)^c = A \cap (\bigcup_{i=1}^{n-1} E_i)$. Now for E_n, a measurable set, we have $\mu^*(A \cap (\bigcup_{i=1}^{n} E_i)) = \mu^*((A \cap (\bigcup_{i=1}^{n} E_i)) \cap E_n) + \mu^*((A \cap (\bigcup_{i=1}^{n} E_i)) \cap E_n^c) = \mu^*(A \cap E_n) + \mu^*(A \cap (\bigcup_{i=1}^{n-1} E_i)) = \mu^*(A \cap E_n) + \sum_{i=1}^{n-1} \mu^*(A \cap E_i)$ by assumption for $n-1$ sets . $\square$

Theorem 8.4. *The class $\mathcal{B}$ of μ^*- measurable sets is a $\sigma-$ algebra. If $\overline{\mu}$ is μ^* restricted to $\mathcal{B}$, then $\overline{\mu}$ is countably additive.*

Note that to prove that a set E is measurable we only need to show that for every A we have $\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$ since μ^* is subadditive.When $\mu^* A = \infty$ the inequality is clearly true, so we only need to show it for sets A with $\mu^* A$ finite; that is $\mu^* A < \infty$.

Proof. First we will show that $\mathcal{B}$ is an algebra of sets. Clearly the empty set is measurable. Now if $E \in \mathcal{B}$ then for each set A, $\mu^* A = \mu^*(A \cap E) = \mu^*(A \cap E^c)$; since this definition is symmetric in E and E^c, E^c is measurable whenever E is. If $E_1, E_2 \in \mathcal{B}$ and $A \cap (E_1 \cup E_2) \in P(X)$, we have $\mu^*(A \cap (E_1 \cup E_2)) = \mu^*((A \cap (E_1 \cup E_2)) \cap E_1) + \mu^*((A \cap (E_1 \cup E_2)) \cap E_1^c)$ $= \mu^*(A \cap E_1) + \mu^*(A \cap (E_2 \cap E_1^c))$. Hence $\mu^*(A \cap E_1) = \mu^*(A \cap (E_1 \cup E_2)) - \mu^*(A \cap (E_2 \cap E_1^c))$.

Now $\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c)$ and by substituting for $\mu^*(A \cap E_1)$ we get $\mu^*(A) = \mu^*(A \cap (E_1 \cup E_2)) - \mu^*(A \cap (E_2 \cap E_1^c)) + \mu^*(A \cap E_1^c)$.
Now we just need $\mu^*(A \cap (E_1 \cup E_2)^c) = \mu^*(A \cap E_1^c) - \mu^*(A \cap (E_2 \cap E_1^c))$ which becomes $\mu^*(A \cap (E_1^c \cap E_2^c)) = \mu^*(A \cap E_1^c) - \mu^*(A \cap (E_2 \cap E_1^c))$.

Moreover $E_2 \in \mathcal{B}$ and $A \cap E_1^c \in P(X)$ imply $\mu^*(A \cap E_1^c) = \mu^*((A \cap E_1^c) \cap E_2) +$ $\mu^*((A \cap E_1^c) \cap E_2^c)$ which gives us what we needed. Hence $\mu^*(A) = \mu^*(A \cap (E_1 \cup E_2)) +$ $\mu^*(A \cap (E_1 \cup E_2)^c)$; that is $E_1 \cup E_2 \in \mathcal{B}$. We have shown that $\mathcal{B}$ is an algebra.

Now we will show that $\mathcal{B}$ is a sigma algebra. Let $\{E_i\}$ be a sequence of measurable sets. Need to show $E = \bigcup_i^\infty E_i$ is also measurable. Without loss of generality we may assume $E_i \cap E_j = \emptyset$ when $i \neq j$. We know $\mu^*(A) = \mu^*(A \cap (\bigcup_{i=1}^n E_i)) + \mu^*(A \cap (\bigcup_{i=1}^n E_i)^c)$ for all n since $\mathcal{B}$ is an algebra.

Then by lemma 8.3, $\mu^*(A) = \sum_{i=1}^n \mu^*(A \cap E_i) + \mu^*(A \cap (\bigcup_{i=1}^n E_i)^c) \geq \sum_{i=1}^n \mu^*(A \cap E_i) + \mu^*(A \cap (\bigcup_{i=1}^\infty E_i)^c)$ for all n; hence $\mu^*(A) \geq \sum_{i=1}^\infty \mu^*(A \cap E_i) + \mu^*(A \cap (\bigcup_{i=1}^\infty E_i)^c)$. Now $\sum_{i=1}^\infty \mu^*(A \cap E_i) \geq \mu^*(A \cap (\bigcup_{i=1}^\infty E_i))$ by subadditivity; therefore $\mu^*(A) \geq \mu^*(A \cap (\bigcup_{i=1}^\infty E_i)) + \mu^*(A \cap (\bigcup_{i=1}^\infty E_i)^c)$. Also by subadditivity of μ^* we have the inequality in the other direction. Hence, we must have $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$.

Next we will show that $\overline{\mu}$ is countably additive for a sequence $\{E_n\} \in \mathcal{B}$ of pairwise disjoint sets, that is $\overline{\mu}(\cup E_i) = \sum \overline{\mu} E_i$, where $\overline{\mu}$ is μ^* restricted to $\mathcal{B}$.

When measure is infinite, $\overline{\mu}(\cup E_i) \geq \sum \overline{\mu}(E_i)$ since $\overline{\mu}(\cup E_i) = \infty \geq \sum \overline{\mu}(E_i)$. Now when the measure is finite, we will first show finite additivity of $\overline{\mu}$. Let E_1 and E_2 be disjoint measurable sets. Since E_1 is measurable, we have $\overline{\mu}(E_1 \cup E_2) = \overline{\mu}((E_1 \cup E_2) \cap$

$E_1) + \overline{\mu}((E_1 \cup E_2) \cap E_1^c) = \overline{\mu}(E_1) + \overline{\mu}(E_2)$. It follows that $\overline{\mu}(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \overline{\mu}(E_i)$ by induction. Hence $\overline{\mu}$ is finitely additive.

Now $\overline{\mu}(\bigcup_{i=1}^\infty E_i) = \overline{\mu}(\bigcup_{i=1}^n E_i) + \overline{\mu}(\bigcup_{i=n+1}^\infty E_i)$. By finite additivity of the outer measure, we have $\overline{\mu}(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^n \overline{\mu}(E_i) + \overline{\mu}(\bigcup_{i=n+1}^\infty E_i)$. This implies, $\overline{\mu}(\bigcup_{i=1}^\infty E_i) \geq \sum_{i=1}^n \overline{\mu}(E_i)$ for all n; hence $\overline{\mu}(\bigcup_{i=1}^\infty E_i) \geq \sum_{i=1}^\infty \overline{\mu}(E_i)$. Now by countable subadditivity we have $\overline{\mu}(\bigcup_{i=1}^\infty E_i) \leq \sum_{i=1}^\infty \overline{\mu}(E_i)$. Hence $\overline{\mu}(\bigcup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \overline{\mu}(E_i)$. □

8.2 From a Measure on an Algebra to a Measure on a σ-algebra Containing the Algebra

Definition 8.5. *By a measure on an algebra we mean a nonnegative extended real-valued set function μ defined on an algebra $\mathcal{A}$ of sets such that:*
i. $\mu(\emptyset) = 0$.
ii. If $\{A_i\}$ is a disjoint sequence of sets in $\mathcal{A}$ whose union is also in $\mathcal{A}$,
then $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu A_i$.
iii. If $\{A_i\}$ is a disjoint sequence of sets in $\mathcal{A}$ then $\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu A_i$ follows from (ii) above.

We will use the measure on the algebra to construct an outer measure μ^*, then by filtering (i.e. applying the Caratheodory's test) we will get the μ^*-measurable sets. This is similar to the process in chapter 1 where we started with the lengths of intervals and used this to construct Lebesgue outer measure, and then by using the Caretheodory process we ended up with Lebesgue measurable sets.

Hence we will show that if we start with a measure on an algebra $\mathcal{A}$ of sets , we may extend it to a measure defined on a sigma algebra $\mathcal{B}$ containing $\mathcal{A}$ [Bar66]. By convention unless otherwise stated we assume all sets are contained in X, that is X is the underlying set on which the algebra is defined. Note that $\emptyset$ as well as X are in the algebra.

The following lemma shows that the measure on the algebra is countably subadditive.

Lemma 8.6. *If $A \in \mathcal{A}$ and if $\{A_i\}$ is any sequence of sets in $\mathcal{A}$ such that $A \subset \bigcup_{i=1}^\infty A_i$, then $\mu A \leq \sum_{i=1}^\infty \mu A_i$.*

Proof. Set $B_n = A \cap A_n \cap A^c_{n-1} \ldots A^c_1$. Since each $B_n \subset A$ we have $\bigcup_{n=1}^{\infty} B_n \subset A$. Moreover $B_n \subset A_n$. Now $A \subset \bigcup_{n=1}^{\infty} B_n$ since for each $x \in A$ there exist m such that if $x \in A_m$ and $x \in A^c_n$ for $n < m$, then $x \in B_m$. Hence $A = \bigcup_{n=1}^{\infty} B_n$. Since $A \in \mathcal{A}$, then $\bigcup_{n=1}^{\infty} B_n \in \mathcal{A}$.

Now since the B_n's are disjoint and in the algebra then by property 2 of the measure on an algebra we get $\mu A = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu B_n$. Moreover each $A_n = B_n \cup (B^c_n \cap A_n)$ by construction. Hence B_n and $(B^c_n \cap A_n)$ are in $\mathcal{A}$ and they are disjoint. It follows that $\mu A_n = \mu(B_n \cup (B^c_n \cap A_n)) = \mu B_n + \mu(B^c_n \cap A_n)$ by property 3 of measure on an algebra. Since $\mu(B^c_n \cap A_n) \geq 0$, we have $\mu A_n \geq \mu B_n$ which implies $\sum_{n=1}^{\infty} \mu A_n \geq \sum_{n=1}^{\infty} \mu B_n = \mu A$. □

Definition 8.7. *We define $\mu^* E = inf \sum_{i=1}^{\infty} \mu A_i$, where $\{A_i\}$ ranges over all sequences from $\mathcal{A}$, where $\mathcal{A}$ is an algebra of sets, such that $E \subset \bigcup_{i=1}^{\infty} A_i$.*

The next corollary shows that μ^* agrees with the measure on the elements in the algebra.

Corollary 8.8. *If $A \in \mathcal{A}$, $\mu^* A = \mu A$.*

Proof. Let $A \in \mathcal{A}$ and $A \subset \bigcup_{i=1}^{\infty} A_i$, where $\{A_i\}$ is any sequence of sets is $\mathcal{A}$, then $\mu A \leq \sum_{i=1}^{\infty} \mu A_i$ by lemma 8.6. Moreover, we can create a disjoint sequence of sets in $\mathcal{A}$ such that $\bigcup_{n=1}^{\infty} B_n = A$. Hence $\mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu B_n = \mu A \leq \bigcup_{i=1}^{\infty} \mu A_i$ for all sequences from $\mathcal{A}$ such that $A \subset \bigcup_{i=1}^{\infty} A_i$. It follows $\sum_{n=1}^{\infty} \mu B_n = inf \sum_{i=1}^{\infty} \mu A_i = \mu^* A$ that is $\mu A = \mu^* A$.□

In the next lemma, we will verify that the extension μ^* of μ is in fact an outer measure, as described in definition 8.1.

Lemma 8.9. *The set function μ^* is an outer measure.*

Proof. If $\emptyset \in \mathcal{A}$, then $\mu^*(\emptyset) = \mu(\emptyset) = 0$ by corollary 8.8. Let $A \subset B$ then for all $\epsilon > 0$ there exist $B \subset \bigcup_{i=1}^{\infty} A_i$ with $A_i \in \mathcal{A}$ such that $\sum_{i=1}^{\infty} \mu(A_i) \leq \mu^*(B) + \epsilon$. Moreover $A \subset B \subset \bigcup_{i=1}^{\infty} A_i$, implies $\mu^*(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$. Hence $\mu^*(A) \leq \mu^*(B) + \epsilon$ for all $\epsilon > 0$. Therefore $\mu^*(A) \leq \mu^*(B)$.

Next we will show that μ^* is countably subadditive. Suppose $E \subset \bigcup_{i=1}^{\infty} E_i$.

Case 1: If $\sum_{i=1}^{\infty}\mu^*(E_i)=\infty$ then $\mu^*(E_i)\leq\sum_{i=1}^{\infty}\mu^*(E_i)=\infty$.

Case 2: If $\sum_{i=1}^{\infty}\mu^*(E_i)<\infty$ then for all $\epsilon>0$ there exist $\{A_{in}\}$ such that $E_i\subset\bigcup_{n=1}^{\infty}A_{in}$ with $A_{in}\in\mathcal{A}$ and $\sum_{n=1}^{\infty}\mu^*(A_{in})\leq\mu^*(E_i)+\frac{\epsilon}{2^i}$.

By design $E\subset\bigcup_{i=1}^{\infty}E_i\subset\bigcup_{i=1}^{\infty}\left(\bigcup_{n=1}^{\infty}A_{in}\right)$.

Hence $\mu^*(E)\leq\sum_{i=1}^{\infty}\left(\sum_{n=1}^{\infty}\mu(A_{in})\right)\leq\sum_{i=1}^{\infty}(\mu^*(E_i)+\frac{\epsilon}{2^i})=\sum_{i=1}^{\infty}\mu^*(E_i)+\sum_{i=1}^{\infty}\frac{\epsilon}{2^i}=\sum_{i=1}^{\infty}\mu^*(E_i)+\epsilon$ since $\epsilon\sum_{i=1}^{\infty}\frac{1}{2^i}=\epsilon$.

Since $\mu^*(E)\leq\sum_{i=1}^{\infty}\mu^*(E_i)+\epsilon$ for all $\epsilon>0$, we have $\mu^*(E)\leq\sum_{i=1}^{\infty}\mu^*(E_i)$. □

Lemma 8.10. *If $A\in\mathcal{A}$, then A is measurable with respect to μ^*.*

Proof. Let E be an element of the power set of X.

For all $\epsilon>0$ there exist $\{A_i\}$ with $A_i\in\mathcal{A}$ such that $E\subset\bigcup_{i=1}^{\infty}A_i$ and $\mu^*(E)+\epsilon\geq\sum_{i=1}^{\infty}\mu A_i$. Note $A_i=(A_i\cap A)\cup(A_i\cap A^c)$ implies $\mu(A_i)=\mu(A_i\cap A)+\mu(A_i\cap A^c)$ by finite additivity of μ. Now $\sum_{i=1}^{\infty}\mu(A_i)=\sum_{i=1}^{\infty}\mu(A_i\cap A)+\sum_{i=1}^{\infty}\mu(A_i\cap A^c)$. Moreover $\sum_{i=1}^{\infty}\mu^*(A_i\cap A)\geq\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A)\right)$ and $\sum_{i=1}^{\infty}\mu^*(A_i\cap A^c)\geq\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A^c)\right)$ by subadditivity of μ^*.

In particular, $E\cap A\subset\bigcup_{i=1}^{\infty}(A_i\cap A)$ and $E\cap A^c\subset\bigcup_{i=1}^{\infty}(A_i\cap A^c)$ imply $\mu^*(E\cap A)\leq\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A)\right)$ and $\mu^*(E\cap A^c)\leq\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A^c)\right)$. It follows $\mu^*(E\cap A)+\mu^*(E\cap A^c)\leq\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A)\right)+\mu^*\left(\bigcup_{i=1}^{\infty}(A_i\cap A^c)\right)\leq\sum_{i=1}^{\infty}\mu^*(A_i\cap A)+\sum_{i=1}^{\infty}\mu^*(A_i\cap A^c)=\sum_{i=1}^{\infty}\mu A_i$. Now $\mu^*(E)+\epsilon\geq\sum_{i=1}^{\infty}\mu A_i$; hence, $\mu^*E+\epsilon\geq\mu^*(E\cap A)+\mu^*(E\cap A^c)$ for all $\epsilon>0$. It follows $\mu^*E\geq\mu^*(E\cap A)+\mu^*(E\cap A^c)$. Moreover, $\mu^*E\leq\mu^*(E\cap A)+\mu^*(E\cap A^c)$ by subadditivity of μ^*. Therefore $\mu^*E=\mu^*(E\cap A)+\mu^*(E\cap A^c)$.
□

Theorem 8.11. *(Caratheodory) Let μ be a measure on an algebra $\mathcal{A}$ and μ^* the outer measure induced by μ. Then the restriction $\overline{\mu}$ of μ^* to the μ^*-measurable sets is an extension of μ to a σ-algebra containing $\mathcal{A}$.*

Proof. Let μ be a measure on an algebra $\mathcal{A}$, then the induced outer measure μ^*, on a set E, is defined as follows $\mu^*E=inf\sum_{i=1}^{\infty}\mu A_i$, where $\{A_i\}$ ranges over all sequences from $\mathcal{A}$, such that $E\subset\bigcup_{i=1}^{\infty}A_i$.
Now define a set E to be measurable with respect to μ^* if for every set A we have $\mu^*(A)=\mu^*(A\cap E)+\mu^*(A\cap E^c)$.

By theorem 8.4 the class $\mathcal{B}$ of μ^*-measurable sets is a σ-algebra, and when $\overline{\mu}$ is μ^* restricted to $\mathcal{B}$ then $\overline{\mu}$ is countably additive.

Now by corollary 8.8, we have that if $A \in \mathcal{A}$, then $\mu^* A = \mu A$; that is, the outer measure is an extension of the measure on the algebra.

Also by lemma 8.10, we have that if $A \in \mathcal{A}$, then A is μ^* measurable; that is, everything in the algebra is measurable. Hence the $\sigma-$ algebra contains the algebra. So we have $\overline{\mu}$ is an extension of μ to a $\sigma-$ algebra containing $\mathcal{A}$. □

8.3 From a Semi-algebra to a Measure on an Algebra

Definition 8.12. *We say that a collection C of subsets of X is a semi-algebra of sets if the intersection of any two sets in C is again in C and the complement of any set in C is a finite disjoint union of sets in C.*

In this section we will see that if we start with a semi-algebra and a nonnegative set function defined on the semi-algebra then we can extend this set function to a measure on the algebra generated by the semi-algebra [Roy88].

Proposition 8.13. *Let C be a semi-algebra of sets and μ a nonnegative set function defined on C with $\mu(\emptyset) = 0$. Then μ has a unique extension to a measure on the algebra $\mathcal{A}$ generated by C if the following conditions are satisfied:*
i. If a set c in C is the union of a finite disjoint collection $\{c_i\}$ of the sets in C, then $\mu c = \sum \mu c_i$.
ii. If a set c in C is the union of a countable disjoint collection $\{c_i\}$ of sets in C, then $\mu c \leq \sum \mu c_i$.

To prove proposition 8.13, we will start by defining $\widetilde{\mu}(A) = \sum_{i=1}^{n} \mu(c_i)$ where $A = \bigcup_{i=1}^{n} c_i^{disj}$ with $c_i \in C$. We will show that $\widetilde{\mu}$ is a measure on the algebra generated by the semi-algebra, i.e. $\widetilde{\mu}$ is an extension of μ on the algebra generated by the semi-algebra. We begin by showing that $\widetilde{\mu}$ is a well defined function on $\mathcal{A}$ (lemma 8.14 and 8.15).

Lemma 8.14. *If $A = \bigcup_{i=1}^{n} c_i^{disj} = \bigcup_{j=1}^{m} d_j^{disj}$ with $c_i, d_j \in C$ then $\sum_{i=1}^{n} \mu(c_i) = \sum_{j=1}^{m} \mu(d_j)$.*

Proof. Note $c_k = \left(\bigcup_{i=1}^{n} c_i\right) \cap c_k = \left(\bigcup_{j=1}^{m} d_j\right) \cap c_k = \bigcup_{j=1}^{m} (d_j \cap c_k)$.

In particular, $(d_j \cap c_k) \in C$ for each j implies, by (i) of proposition 8.13, that $\mu(c_k) = \sum_{j=1}^{m} \mu(d_j \cap c_k)$. Similarly one can show $\mu(d_j) = \sum_{i=1}^{n} (c_i \cap d_j)$. Hence, $\sum_{i=1}^{n} \mu(c_i) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \mu(d_j \cap c_i)\right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \mu(d_j \cap c_i)\right) = \sum_{j=1}^{m} \mu(d_j)$. □

Lemma 8.15. *If* $\bigcup_{j=1}^{m} d_j^{disj} = \bigcup_{i=1}^{\infty} c_i^{disj}$ *with* $c_i, d_j \in \mathcal{C}$ *then* $\sum_{j=1}^{m} \mu(d_j) = \sum_{i=1}^{\infty} \mu(c_i)$.

Proof. $d_j = \bigcup_{i=1}^{\infty}(c_i \cap d_j)$ implies, by (ii) of proposition 8.13, that $\mu(d_j) \leq \sum_{i=1}^{\infty} \mu(c_i \cap d_j)$. It follows $\sum_{j=1}^{m} \mu(d_j) \leq \sum_{j=1}^{m} \left(\sum_{i=1}^{\infty} \mu(c_i \cap d_j)\right)$.

In particular, $\bigcup_{j=1}^{m}(c_i \cap d_j) = c_i$ implies, by (i) of proposition 8.13, that $\sum_{j=1}^{m} \mu(c_i \cap d_j) = \mu(c_i)$. Hence we must have $\sum_{j=1}^{m} \mu(d_j) \leq \sum_{i=1}^{\infty} \mu(c_i)$.

Moreover, $\left(\bigcup_{j=1}^{m} d_j - \bigcup_{i=1}^{n} c_i\right) \cup (\bigcup_{i=1}^{n} c_i) = \left(\bigcup_{j=1}^{m}(d_j \cap (\bigcap_{i=1}^{n} c_i^c))\right) \cup (\bigcup_{i=1}^{n} c_i)$ $= \bigcup_{j=1}^{m} d_j$ implies $d_k = \left(\bigcup_{j=1}^{m}(d_j \cap (\bigcap_{i=1}^{n} c_i^c) \cap d_k)\right) \cup (\bigcup_{i=1}^{n} c_i \cap d_k)$. Now by property (ii) $\mu(d_k) \geq \sum_{i=1}^{n} \mu(c_i \cap d_k)$ for all n. Hence, $\sum_{k=1}^{m} \mu(d_k) \geq \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \mu(c_i \cap d_k)\right)$ $= \sum_{i=1}^{n} \left(\sum_{k=1}^{m} \mu(c_i \cap d_k)\right) = \sum_{i=1}^{n} \mu(c_i)$ since $\mu(c_i) = \sum_{k=1}^{m} \mu(c_j \cap d_k)$. Therefore, $\sum_{k=1}^{m} \mu(d_k) \geq \sum_{i=1}^{\infty} \mu(c_i)$ which implies $\sum_{j=1}^{m} \mu(d_k) = \sum_{i=1}^{\infty} \mu(c_i)$. □

In lemma 8.16 we will show everything in $\mathcal{A}$ is a finite union of disjoint elements of $\mathcal{C}$.

Lemma 8.16. *If* $A = \bigcup_{i=1}^{n} c_i$ *where* $c_i \in \mathcal{C}$ *then* $A = \bigcup_{j=1}^{m} d_j^{disj}$ *where* $d_j \in \mathcal{C}$.

Proof. Let $A = \bigcup_{i=1}^{n} c_i$. Define $a_1 = c_1$, $a_2 = c_2 \bigcap c_1^c$, $a_3 = c_3 \bigcap c_2^c \bigcap c_1^c$ then it follows $a_n = c_n \bigcap c_{n-1}^c \cdots \bigcap c_1^c$. By design $A = \bigcup_{i=1}^{n} a_i^{disj}$. Since $\mathcal{C}$ is a semi-algebra, for each i we have $a_i = c_i \bigcap c_{i-1}^c \cdots \bigcap c_1^c = \bigcup_{j=1}^{k_i} b_{ij}^{disj}$ and $b_{ij} \in \mathcal{C}$. Hence $A = \bigcup_{i=1}^{n}(\bigcup_{j=1}^{k_i} b_{ij}^{disj})$. □

Next we will show that if a countable union of disjoint sets in the algebra happen to be in the algebra, then $\widetilde{\mu}$ is countably additive.

Theorem 8.17. *If* $A = \bigcup_{j=1}^{\infty} A_j^{disj}$ *with* $A, A_j \in \mathcal{A}$ *then* $\widetilde{\mu}A = \sum_{j=1}^{\infty} \widetilde{\mu}A_j$.

Proof. $A = \bigcup_{i=1}^{k} c_i^{disj}$ where $c_i \in \mathcal{C}$ and $A_j = \bigcup_{i=1}^{k_j} d_{ji}^{disj}$ where $d_{ji} \in \mathcal{C}$ by lemma 8.16. In particular, $\bigcup_{i=1}^{k} c_i^{disj} = \bigcup_{j=1}^{\infty} \left(\bigcup_{i=1}^{k_j} d_{ji}^{disj}\right)$ which implies by lemma 8.15 that $\widetilde{\mu}A = \sum_{i=1}^{k} \mu(c_i) = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{k_j} \mu(d_{ji})\right) = \sum_{j=1}^{\infty} \widetilde{\mu}A_j$. □

Corollary 8.18. *If* $A = \bigcup_{i=1}^{n} A_i^{disj}$ *with* $A_i \in \mathcal{A}$ *then* $\widetilde{\mu}A = \sum_{i=1}^{n} \widetilde{\mu}A_i$.

Proof. $A = \bigcup_{i=1}^{\infty} A_i^{disj}$ with $A_m = \emptyset$ for $m \geq n$. Hence by theorem 8.17 $\widetilde{\mu}A = \sum_{j=1}^{\infty} \widetilde{\mu}A_j = \sum_{j=1}^{n} \widetilde{\mu}A_j$ since $\widetilde{\mu}(\emptyset) = 0$. □

We have thus shown that $\widetilde{\mu}$ is a measure on $\mathcal{A}$.

An example of a semialgebra are the intervals. The length of the intervals is a nonnegative set function that meets the criteria of proposition 8.13. Hence we can extend the length function to a measure on the algebra generated by the intervals. Moreover, by the Caratheodory process we can extend the measure on the algebra to a sigma algebra containing the algebra which is countably additive. We call the elements of this sigma algebra Lebesgue measurable sets, and the measure on this sigma algebra is called Lebesgue measure. By design the Lebesgue measure of an interval will be its length.

Chapter 9

Conclusion

The procedure that Lebesgue used to define his Lebesgue measure can be generalized to create new measures. Starting with a semialgebra $\mathcal{C}$ defined on a set X, we can define a measure μ on $\mathcal{C}$. We can extend this measure to the algebra generated by $\mathcal{C}$. Then we can define an outer measure μ^* with respect to $\mathcal{C}$. Now, everything in the power set of X has an outer measure. Since we want to have countable additivity, we filter the sets in the power set of X by applying the Caratheodory process. The sets obtained after filtering form a sigma algebra which contains the semialgebra and the algebra. Moreover the outer measure restricted to this sigma algebra is countably additive. Once a measure has been obtained we can define a new theory of integration based on that measure. It is interesting to note that one of the new integrals developed called the Henstok integral bases its theory of integration on Riemann instead of Lebesgue.

Bibliography

[Bar66] Robert G. Bartle. *The elements of Integration.* John Wiley and Sons, Inc., New York, 1966.

[Bur98] Frank Burk. *Lebesgue Measure and Integration: An Introduction.* John Wiley and Sons, Inc., New York, 1998.

[Mun00] James R. Munkres. *Topology.* Prentice Hall, Inc., New Jersey, 2000.

[Roy88] H. L. Royden. *Real Analysis.* The Macmillan Company, New York, third edition, 1988.

[Rud87] Walter Rudin. *Real Complex Analysis.* McGraw-Hill, Inc., California, third edition, 1987.

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