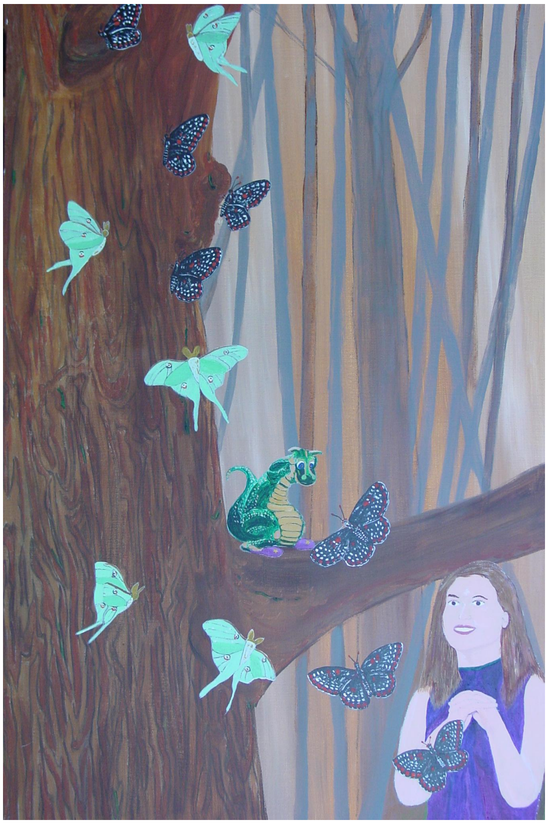
Measure and Integration: First Steps



We are all made of the same stuff, Dewey.

Jim Peterson
Department of Mathematical Sciences
Clemson University
email: petersj@clemson.edu

December 12, 2006

Based On:

Handwritten Notes For Measure and Integration

MTHSC 822

1995 - 2006

Jim has been steadily shrinking since 1996 and is currently only able to be seen when standing on a large box. This has forced him to alter his lecture style somewhat so that his students can see him and he is considering the use of platform shoes. However, if he can ever remember the code to his orbiting spaceship, he will be able to access the shape changing facilities on board and solve the problem permanently. However, he does what he can with this handicap.

There were many problems with the handwritten version of these notes, but with the help of my many students in this class over the years, they have gotten better. However, being written in my handwriting was not a good thing. In fact, one student some years ago was so unhappy with the handwritten notes, that they sent me the names of four engineering professors who actually typed their class notes nicely. He told me I could learn from them how to be a better teacher. Well, it has taken me awhile, but at least the notes are being typed. The teaching part, though, is another matter.

I wish to thank all my students for helping me by listening to what I say in my lectures, finding my typographical errors and my other mistakes. I am always hopeful that my efforts help my students to some extent and also impart some of my enthusiasm for the subject. Of course, the reality is that I have been damaging students for years by forcing them to learn these abstract things. This is why I tell people at parties I am a roofer or electrician. If I am identified as a mathematician, it could go badly given the terrible things I inflict on the students in my classes. Who knows whom they have told about my tortuous methods. Hence, anonymity is best.

Still, by writing these notes, I have gone public. Sigh. This could be bad. So before I am taken out with a very public hit, I, of course, want to thank my family for all their help in many small and big ways. It is amazing to me that I have been teaching this material since my children were in grade school and now my youngest is getting ready to begin college. They have put up with my vacant stares way too much.

Contents

Contents i

I Introductory Matter 1

1 Introduction 3
1.1 Senior Level Analysis 4
1.2 The Graduate Analysis Courses 5
1.3 More Advanced Courses 6
1.4 Teaching The Measure and Integration Course 7

II The Main Event 9

2 An Overview Of Riemann Integration 11
2.1 Integration 11
2.1.1 A Riemann Sum Example 13
2.1.2 The Riemann Integral As A Limit 14
2.1.3 The Fundamental Theorem Of Calculus 16
2.1.4 The Cauchy Fundamental Theorem Of Calculus 19
2.1.5 Applications 21
2.1.6 Simple Substitution Techniques 22
2.2 The Riemann Integral of Functions With Jumps 25
2.2.1 Removable Discontinuity 26
2.2.2 Jump Discontinuity 27
2.2.3 Homework 29

3 Functions Of Bounded Variation 31
3.1 Partitions 32

3.1.1 Homework . . . 33
3.2 Monotone Functions . . . 33
3.2.1 Worked Out Example . . . 41
3.2.2 Homework . . . 43
3.3 Functions of Bounded Variation . . . 44
3.3.1 Homework . . . 49
3.4 The Total Variation Function . . . 49
3.5 Continuous Functions of Bounded Variation . . . 52

4 The Theory Of Riemann Integration 55
4.1 Defining The Riemann Integral . . . 55
4.2 The Existence of the Riemann Integral: Darboux Integration . . . 58
4.3 Properties Of The Riemann Integral . . . 65
4.4 What Functions Are Riemann Integrable? . . . 69
4.5 Further Properties of the Riemann Integral . . . 71
4.6 The Fundamental Theorem Of Calculus . . . 74
4.6.1 Homework . . . 80
4.7 Substitution Type Results . . . 80
4.8 When Do Two Functions Have The Same Integral? . . . 83

5 Further Riemann Integration Results 87
5.1 The Limit Interchange Theorem for Riemann Integration . . . 87
5.2 Showing Functions Are Riemann Integrable . . . 92
5.3 Sets Of Content Zero . . . 94

6 Cantor Set Experiments 101
6.1 The Generalized Cantor Set . . . 101
6.2 Representing The Generalized Cantor Set . . . 104
6.3 The Cantor Function . . . 106

7 The Riemann-Stieltjes Integral 109
7.1 Standard Properties Of The Riemann - Stieljes Integral . . . 110
7.2 Step Functions As Integrators . . . 112
7.3 Monotone Integrators . . . 117
7.4 The Riemann - Stieljes Equivalence Theorem . . . 118
7.5 Properties Of The Riemann Integral . . . 119
7.6 Bounded Variation Integrators . . . 121

8 Further Riemann Stieljes Results 125
8.1 The Riemann - Stieljes Fundamental Theorem Of Calculus 125
8.2 Existence Results 128
8.3 Worked Out Examples Of Riemann Stieljes Computations 131
8.4 Homework 136

9 Measurable Functions and Spaces 139
9.1 Examples 140
9.2 The Borel Sigma - Algebra of $\Re$ 141
9.2.1 Homework 143
9.3 The Extended Borel Sigma Algebra 143
9.4 Measurable Functions 145
9.4.1 Examples 147
9.5 Properties of Measurable Functions 148
9.6 Extended Valued Measurable Functions 150
9.7 Homework 157

10 Measure And Integration 159
10.1 Some Basic Properties Of Measures 161
10.2 Integration 166
10.3 Complete Measures And Equality a.e. 171
10.4 Convergence Theorems 174
10.5 Extending Integration To Extended Real Valued Functions 180
10.6 Properties Of Summable Functions 184
10.7 The Dominated Convergence Theorem 186
10.8 Homework 189

11 The $\mathcal{L}_p$ Spaces 191
11.1 The General L_p spaces 195
11.2 The World Of Counting Measure 204
11.3 Equivalence Classes of Essentially Bounded Functions 206
11.4 The Hilbert Space L_2 212
11.5 Homework 213

12 Constructing Measures 215
12.1 Measures From Outer Measures 215
12.2 Measures From Metric Outer Measures 222
12.3 Constructing Outer Measures 227
12.4 Worked Out Problems 231
12.5 Homework 233

13 Lebesgue Measure **235**
13.1 Lebesgue Outer Measure and Measure . . . 235
13.2 Lebesgue Outer Measure Is A Metric Outer Measure . . . 245
13.3 Lebesgue - Stieljes Outer Measure and Measure . . . 250
13.4 Homework . . . 255

14 Modes Of Convergence **257**
14.1 Subsequence Extraction . . . 259
14.2 Egoroff's Theorem . . . 267
14.3 Vitali Convergence Theorem . . . 269
14.4 Summary . . . 274
14.5 Homework . . . 277

15 Decomposition Of Measures **279**
15.1 Basic Decomposition Results . . . 279
15.2 The Variation Of A Charge . . . 286
15.3 Absolute Continuity Of Charges . . . 289
15.4 The Radon - Nikodym Theorem . . . 291
15.5 The Lebesgue Decomposition of a Measure . . . 297
15.6 Homework . . . 301

16 Connections To Riemann Integration **303**

17 Fubini Type Results **305**

18 Interesting Questions **307**
18.1 Midterm Examination . . . 307
18.2 Final Examination . . . 309

III References **313**

Bibliography **315**

IV Detailed Indices **317**

Index **319**

Part I

Introductory Matter

Chapter 1

Introduction

We believe that all students who are seriously interested in mathematics at the Master's and Doctoral level should have a passion for analysis even if it is not the primary focus of their own research interests. So you should all understand that my own passion for the subject will shine though in the notes that follow! And, it goes without saying that we assume that you are all mature mathematically and eager and interested in the material! Now, the present course focuses on the topics of *Measure* and *Integration* from a very abstract point of view, but it is very helpful to place this course into its proper context. Also, for those of you who are preparing to take the qualifying examination in in analysis, the overview below will help you see why all this material fits together into a very interesting web of ideas. These ideas are covered in the we will discuss below. In outline form, these courses would cover the following material using textbooks equivalent to the ones listed below:

(A): Undergraduate Analysis, text **Advanced Calculus: An Introduction to Analysis**, by Watson Fulks. Here these are MTHSC 453 and MTHSC 454.

(B): Introduction to Abstract Spaces, text **Introduction to Functional Analysis and Applications**, by Ervin kreyszig. Here this is MTHSC 821.

(C): Measure Theory and Abstract Integration, texts **General Theory of Functions and Integration**, by Angus Taylor and **Real Analysis**, by Royden. Here this is MTHSC 822.

In addition, a nice book that organizes the many interesting examples and counterexamples in this area is a nice one to have on your shelf. We recommend the text **Counterexamples in Analysis** by Gelbaum and Olmstead. There are thus essentially five courses required to teach you enough of the concepts of mathematical analysis to enable you to read technical literature (such as engineering, control, physics, mathematics, statistics and so forth) at the beginning research level. Here are some more details about these courses.

1.1 Senior Level Analysis

Typically, this is a full two semester sequence that discusses thoroughly what we would call the analysis of functions of a real variable. Here, this is the sequence MTHSC SC 453–454. This two semester sequence covers the following:

Advanced Calculus I: MTHSC 453: This course studies sequences and functions whose domain is simply the real line. There are, of course, many complicated ideas, but everything we do here involves things that act on real numbers to produce real numbers. If we call these things that act on other things, **OPERATORS**, we see that this course is really about real–valued operators on real numbers. This course invests a lot of time in learning how to be precise with the notion of convergence of sequences of objects, that happen to be real numbers, to other numbers.

1. Basic Logic, Inequalities for Real Numbers, Functions
2. Sequences of Real Numbers, Convergence of Sequences
3. Subsequences and the Bolzano–Weierstrass Theorem
4. Cauchy Sequences
5. Continuity of Functions
6. Consequences of Continuity
7. Uniform Continuity
8. Differentiability of Functions
9. Consequences of Differentiability
10. Taylor Series Approximations

Advanced Calculus II: MTHSC 454: In this course, we rapidly become more abstract. First, we develop carefully the concept of the Riemann Integral. We show that although differentiation is intellectually quite a different type of limit process, it is intimately connected with the Riemann integral. Also, for the first time, we begin to explore the idea that we could have sequences of objects other than real numbers. We study carefully their convergence properties. We learn about two fundamental concepts: pointwise and uniform convergence of sequences of objects called functions. We are beginning to see the need to think about sets of objects, such as functions, and how to define the notions of convergence and so forth in this setting.

1. The Riemann Integral
2. Sequences of Functions
3. Uniform Convergence of Sequence of Functions
4. Series of Functions

1.2 The Graduate Analysis Courses

There are three basic courses here. First, linear analysis (MTHSC 821), then measure and integration (MTHSC 822) and finally, functional analysis (MTHSC 927).

Introductory Linear Analysis: MTHSC 821: We now begin to rephrase all of our knowledge about convergence of sequence of objects in a much more general setting.

1. Metric Spaces: A set of objects and a way of measuring distance between objects which satisfies certain special properties. This function is called a **metric** and its properties were chosen to mimic the properties that the absolute value function has on the real line. We learn to understand convergence of objects in a general metric space. It is really important to note that there is NO additional structure imposed on this set of objects; no linear structure (i.e. vector space structure), no notion of a special set of elements called a basis which we can use to represent arbitrary elements of the set. The **metric** in a sense generalizes the notion of distance between numbers. We can't really measure the size of an object by itself, so we do not yet have a way of generalizing the idea of size or length.

   A fundamentally important concept now emerges: the notion of completeness and how it is related to our choice of metric on a set of objects. We learn a clever way of constructing an abstract representation of the completion of any metric space, but at this time, we have no practical way of seeing this representation.

2. Normed Spaces: We add linear structure to the set of objects and a way of measuring the magnitude of an object; that is, there is now an operation we think of as addition and another operation which allows us to scale objects and a special function called a **norm** whose value for a given object can be thought of as the object's magnitude. We then develop what we mean by convergence in this setting. Since we have a vector space structure, we can now begin to talk about a special subset of objects called a **basis** which can be used to find a useful way of representing an arbitrary object in the space.

   Another most important concept now emerges: the cardinality of this basis may be finite or infinite. We begin to explore the consequences of a space being finite versus infinite dimensional.

3. Inner Product Spaces: To a set of objects with vector space structure, we add a function called an **inner product** which generalizes the notion of dot product of vectors. This has the extremely important consequence of allowing the inner product of two objects to zero even though the objects are not the same. Hence, we can develop an abstract notion of the orthogonality of two objects. This leads to the idea of a basis for the set of objects in which all the elements are mutually orthogonal. We then finally can learn how to build representations of arbitrary objects efficiently.

4. Completions: We learn how to complete an arbitrary metric, normed or inner product space in an abstract way, but we know very little about the practical representations of such completions.

5. Linear Operators: We study a little about functions whose domain is one set of objects and whose range is another. These functions are typically called operators. We learn a little about them here.

6. Linear Functionals: We begin to learn the special role that real-valued functions acting on objects play in analysis. These types of functions are called linear functionals and learning how to characterize them is the first step in learning how to use them. We just barely begin to learn about this here.

Measure Theory: MTHSC 822: This course is about generalizing the notion of the length of an interval. We learn how to develop the notion of the length of an arbitrary subset of the real line; This generalization is called a **measure** on the real line. We then extend this notion to other euclidean spaces ($\Re^n$) and finally develop the notion of measures on arbitrary sets of objects. We use these new generalizations of length to carefully develop a corresponding theory of integration.

The set of objects we now work with is a subset of the power set of of a given set S, $P(S)$. The set S could be a set of real numbers, a set of vectors or a set of any objects. A **measure** is a special kind of real–valued function that acts on sets. The set theoretic nature of this function requires many new tools and a new level of abstraction.

It is only when we have these tools at our disposal that we can discuss in a illuminating way the concepts of weak convergence, and convergence in measure. It turns out that we also now have a way of representing the dual spaces of certain classes of functions. This is extremely powerful in applications. So, in this class, we discuss the following:

1. The Riemann Integral
2. Measure in $\Re^n$, integration with respect to the measure
3. Abstract Measures, abstract integration theory
4. Differentiation and Integration

1.3 More Advanced Courses

It is also recommended that at some point you consider taking a course in what is called Functional Analysis. Here that is MTHSC 927. While not part of the qualifying examination, in this course, we can finally develop in a careful way the necessary tools to work with linear operators, weak convergence and so forth. This is a huge area of mathematics, so there are many possible ways to design an introductory course. A typical such course would cover:

1. The Open Mapping and Closed Graph Theorem
2. An Introduction to General Operator Theory
3. An Introduction to the Spectral Theory of Linear Operators; this is the study of the eigenvalues and eigenobjects for a given linear operator–lots of applications here!
4. Some advanced topic using these ideas: possibilities include
   (a) Existence Theory of Boundary Value Problems
   (b) Existence Theory for Integral Equations
   (c) Existence Theory in Control

1.4 Teaching The Measure and Integration Course

So now that you have seen how the analysis courses all fit together, it is time for the main course. So roll up your sleeves and prepare to work! Let's start with a few more details on what this course on *Measure* and *Integration* will cover.

In this course, we assume mathematical maturity and we tend to follow the *The Enthusiastic "maybe I can get them interested anyway" Approach* in lecturing (so, be warned)! It is difficult to decide where to start converge in this course. There is usually a reasonable fraction of you who have never seen an adequate treatment of Riemann Integration. For example, not everyone may have seen the equivalent of MTHSC 454 where Riemann integration is carefully discussed. We therefore have several versions of this course. We have divided the material into *blocks* as follows: We believe there are a lot of advantages in treating integration abstractly. So, if we covered the Lebesgue integral on $\Re$ right away, we can take advantage of a lot of the special structure $\Re$ has which we don't have in general. It is better for long term intellectual development to see measure and integration approached without using such special structure. Also, all of the standard theorems we want to do are just as easy to prove in the abstract setting, so why specialize to $\Re$? So we tend to do abstract measure stuff first. The core material for **Block 1** is as follows:

1. abstract measure $\boldsymbol{\nu}$ on a sigma - algebra $\mathcal{S}$ of subsets of a universe $\boldsymbol{X}$.
2. measurable functions with respect to a measure $\boldsymbol{\nu}$; these are also called *random variables* when $\boldsymbol{\nu}$ is a probability measure.
3. integration $\int f d\nu$
4. convergence results: monotone convergence theorem, dominated convergence theorem etc.

Then we develop the Lebesgue Integral in $\Re^n$ via outer measures as the great example of a nontrivial measure. So **Block 2** of material is thus

1. outer measures in $\Re^n$
2. caratheodory conditions for measurable sets
3. construction of the Lebesgue sigma algebra
4. connections to Borel sets

Along the way, starting from day one, we have a concurrent thread running which concerns the Cantor sets of measure β. We believe there is a lot of value in working out these complicated things as they serve several purposes. First, they are hard but doable for you no matter what your background. Also, they absolutely require an abstract approach. You can't use Matlab to get a good picture of the construction process. So this helps build your intellectual tool set. So in the first month, while learning abstract measure theory, you will also be doing projects on Cantor sets. In the second month, we start you working through the Cantor singular function and the many consequences of that function.

To fill out the course, we pick topics from the following

1. Riemann and Riemann - Stieljes integration. This would go before **Block 1** if we do it. Call it block **Riemann**.

2. Decomposition of measures – I love this material so this is after **Block 2**. Call it block **Decomposition**.
3. Connection to Riemann integration via absolute continuity of functions. this is actually hard stuff and takes about 3 weeks to cover nicely. Call it **Block Riemann and Lebesgue**. If this is done without **Block Riemann**, you have to do a quick review of Riemann stuff so they can follow the proofs.
4. Fubini type theorems. This would go after **Block 2**. Call this Block **Fubini**.
5. Differentiation via the Vitali approach. This is pretty hard too. Call this **Differentiation**.
6. Treatment of the usual L^p spaces. Call this **Block** L^p.
7. More convergence stuff like convergence in measure, L^p convergence implies convergence of a subsequence pointwise etc. These are hard theorems and to do them right requires a lot of time. Call this **More Convergence**.

We have taught this in at least the following ways. And always, lots of homework and projects, as we believe only hands on work really makes this stuff sink in.

Way 1: Block Riemann, Block 1, Block 2 and Block Decomposition.

Way 2: Block 1, Block 2, Block Decomposition and Block Riemann and Lebesgue.

Way 3: Block 1, Block 2, Block Decomposition and Differentiation.

Way 4: Block 1, Block 2, Block L^p, Block More Convergence and Block Decomposition.

Way 5: Block 1, Block 2, Block Fubini, Block More Convergence and Block Decomposition.

So as you can see it will be an interesting ride!

Part II

The Main Event

Chapter 2

An Overview Of Riemann Integration

In this Chapter, we will give you a quick overview of Riemann integration. There are few real proofs but it is useful to have a quick tour before we get on with the job of extending this material to a more abstract setting. Much of this material can be found in a good Calculus book although the more advanced stuff requires that you look at a book on beginning real analysis such as (Fulks (3) 1978) .

2.1 Integration

You should also have been exposed to the idea of the integration of a function f. There are two intellectually separate ideas here:

1. The idea of a *primitive* or *antiderivative* of a function f. This is any function F which is differentiable and satisfies $F'(t) = f(t)$ at all points in the domain of f. Normally, the domain of f is a finite interval of the form $[a, b]$, although it could also be an infinite interval like all of $\Re$ or $[1, \infty)$ and so on. Note that an *antiderivative* does not require any understanding of the process of Riemann integration at all – only what differentiation is!

2. The idea of the Riemann integral of a function. You should have been exposed to this in your first Calculus course and perhaps a bit more rigorously in your undergraduate second semester analysis course.

Let's review what Riemann Integration involves. First, we start with a bounded function f on a finite interval $[a, b]$. This kind of function f need not be continuous! Then select a finite number of points from the interval $[a, b]$, $\{x_0, x_1, , \ldots, x_{n-1}, x_n\}$. We don't know how many points there are, so a different selection from the interval would possibly gives us more or less points. But for convenience, we will just call the last point x_n and the first point x_0. These points are not arbitrary – x_0 is always a, x_n is always b and they are ordered like this:

$$x_0 = a < x_1 < x_2 < \ \ldots \ < x_{n-1} < x_n = b$$

The collection of points from the interval $[a, b]$ is called a Partition of $[a, b]$ and is denoted by some letter – here we will use the letter $\boldsymbol{\pi}$. So if we say $\boldsymbol{\pi}$ is a partition of $[a, b]$, we know it will have $n+1$ points in it, they will be labeled from x_0 to x_n and they will be ordered left to right with strict inequalities. But, we will not know what value the positive integer n actually is. The simplest Partition $\boldsymbol{\pi}$ is the two point partition $\{a, \ b\}$. Note these things also:

1. Each partition of $n + 1$ points determines n subintervals of $[a, b]$

2. The lengths of these subintervals always adds up to the length of $[a, b]$ itself, $b - a$.

3. These subintervals can be represented as

   $$\{[x_0, x_1], \ [x_1, x_2], \ldots, [x_{n-1}, x_n]\}$$

   or more abstractly as $[x_i, x_{i+1}]$ where the index i ranges from 0 to $n - 1$.

4. The length of each subinterval is $x_{i+1} \ - \ x_i$ for the indices i in the range 0 to $n - 1$.

Now from each subinterval $[x_i, x_{i+1}]$ determined by the Partition $\boldsymbol{\pi}$, select any point you want and call it s_i. This will give us the points s_0 from $[x_0, x_1]$, s_1 from $[x_1, x_2]$ and so on up to the last point, s_{n-1} from $[x_{n-1}, x_n]$. At each of these points, we can evaluate the function f to get the value $f(s_j)$. Call these points an **Evaluation Set** for the partition $\boldsymbol{\pi}$. Let's denote such an evaluation set by the letter $\boldsymbol{\sigma}$. Note there are many such evaluation sets that can be chosen from a given partition $\boldsymbol{\pi}$. We will leave it up to you to remember that when we use the symbol $\boldsymbol{\sigma}$, you must remember it is associated with some partition.

If the function f was nice enough to be positive always and continuous, then the product $f(s_i) \times (x_{i+1} \ - \ x_i)$ can be interpreted as the area of a rectangle. Then, if we add up all these rectangle areas we get a sum which is useful enough to be given a special name: the Riemann sum for the function f associated with the Partition $\boldsymbol{\pi}$ and our choice of evaluation set $\boldsymbol{\sigma} \ = \ \{s_0, \ \ldots \ , s_{n-1}\}$. This sum is represented by the symbol $S(f, \boldsymbol{\pi}, \boldsymbol{\sigma})$ where the things inside the parenthesis are there to remind us that this sum depends on our choice of the function f, the partition $\boldsymbol{\pi}$ and the evaluations set $\boldsymbol{\sigma}$. So formally, we have the definition

Definition 2.1.1. Riemann Sum

The Riemann sum for the bounded function f, the partition $\boldsymbol{\pi}$ and the evaluation set $\boldsymbol{\sigma} = \{s_0, \ldots, s_{n-1}\}$ from $\boldsymbol{\pi}\{x_0, x_1, \ldots, x_{n-1}, x_n\}$ is defined by

$$S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \;=\; \sum_{i=0}^{n-1} f(s_i)\,(x_{i+1} - x_i)$$

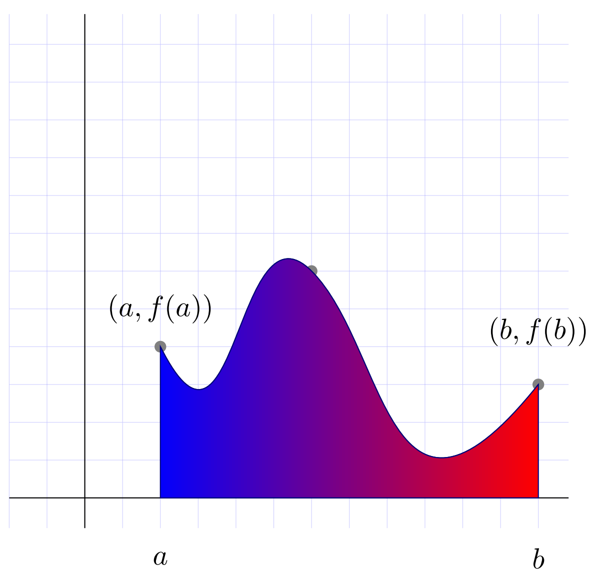
It is pretty misleading to write the Riemann sum this way as it can make us think that the n is always the same when in fact it can change value each time we select a different partition $\boldsymbol{\pi}$. So many of us write the definition this way instead

$$S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \;=\; \sum_{i \,\in\, \boldsymbol{\pi}} f(s_i)\,(x_{i+1} - x_i) = \sum_{\boldsymbol{\pi}} f(s_i)\,(x_{i+1} - x_i)$$

and we just remember that the choice of $\boldsymbol{\pi}$ will determine the size of n.

2.1.1 A Riemann Sum Example

Let's look at an example of all this. In Figure 2.1, we see the graph of a typical function which is always positive on some finite interval $[a, b]$

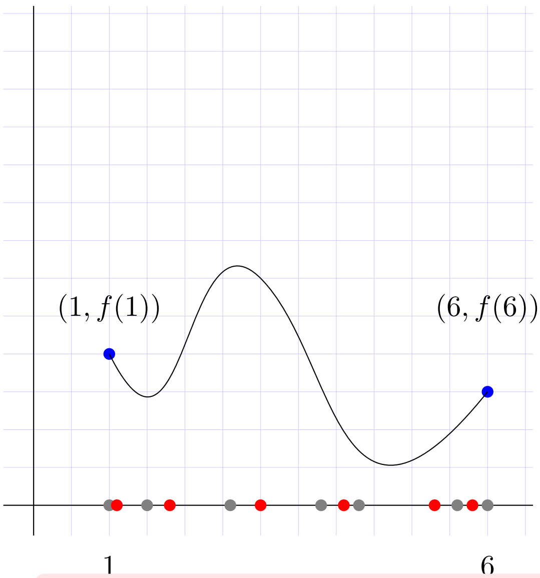


A generic curve f on the interval $[a, b]$ which is always positive. Note the area under this curve is the shaded region.

Figure 2.1: *The Area Under The Curve f*

Next, let's set the interval to be $[1, 6]$ and compute the Riemann Sum for a particular choice of Partition $\boldsymbol{\pi}$ and evaluation set $\boldsymbol{\pi}$. This is shown in Figure 2.2.

We can also interpret the Riemann sum as an approximation to the area under the curve as shown in Figure 2.1. This is shown in Figure 2.3.



The partition is $\boldsymbol{\pi} = \{1.0, 1.5, 2.6, 3.8, 4.3, 5.6, 6.0\}$. Hence, we have subinterval lengths of $x_1 - x_0 = 0.5$, $x_2 - x_1 = 1.1$, $x_3 - x_2 = 1.2$, $x_4 - x_3 = 0.5$, $x_5 - x_4 = 1.3$ and $x_6 - x_5 = 0.4$, giving $\| P \| = 1.3$. Thus,

$$S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) = \sum_{i=0}^{5} f(s_i)\,(x_{i+1} - x_i)$$

For the evaluation set $\boldsymbol{\sigma} = \{1.1, 1.8, 3.0, 4.1, 5.3, 5.8\}$ shown in red in Figure 2.2, we would find the Riemann sum is

$$\begin{aligned} S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) &= f(1.1) \times 0.5 \\ &\quad + f(1.8) \times 1.1 \\ &\quad + f(3.0) \times 1.2 \\ &\quad + f(4.1) \times 0.5 \\ &\quad + f(5.3) \times 1.3 \\ &\quad + f(5.8) \times 0.4 \end{aligned}$$

Of course, since our picture shows a generic f, we can't actually put in the function values $f(s_i)$!

Figure 2.2: *A Simple Riemann Sum*

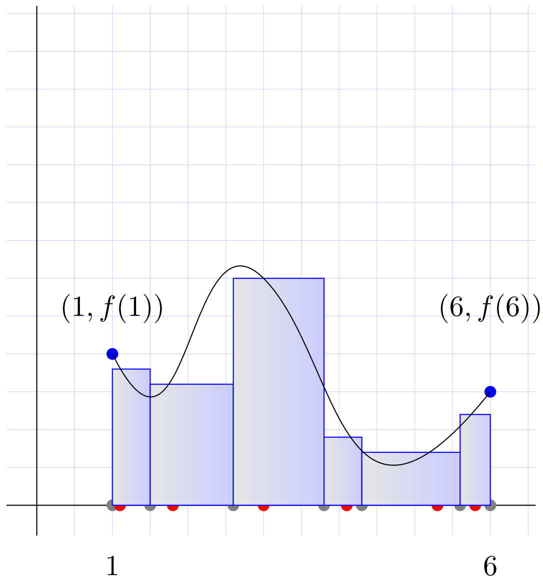
2.1.2 The Riemann Integral As A Limit

We can construct many different Riemann Sums for a given function f. To define the Riemann integral of f, we only need a few more things:

1. Each partition $\boldsymbol{\pi}$ has a maximum subinterval length – let's use the symbol $\| \boldsymbol{\pi} \|$ to denote this length. We read the symbol $\| \boldsymbol{\pi} \|$ as the **norm** or **gauge** of $\boldsymbol{\pi}$.

2. Each partition $\boldsymbol{\pi}$ and evaluation set $\boldsymbol{\sigma}$ determines the number $S(f, \boldsymbol{\pi}, \boldsymbol{\sigma})$ by a simple calculation.

3. So if we took a collection of partitions $\boldsymbol{\pi_1}$, $\boldsymbol{\pi_2}$ and so on with associated evaluation sets $\boldsymbol{\sigma_1}$, $\boldsymbol{\sigma_2}$ etc., we would construct a sequence of real numbers $\{S(f, \boldsymbol{\pi_1}, \boldsymbol{\sigma_1}), S(f, \boldsymbol{\pi_2}, \boldsymbol{\sigma_2}), \ldots, , S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}), \ldots, \}$. Let's assume the norm of the partition $\boldsymbol{\pi_n}$ gets smaller all the time; i.e. $\lim_{n \to \infty} \| \boldsymbol{\pi_n} \| = 0$. We could then ask if this sequence of numbers converges to something.



The partition is $\boldsymbol{\pi} = \{1.0, 1.5, 2.6, 3.8, 4.3, 5.6, 6.0\}$.

Figure 2.3: *The Riemann Sum As An Approximate Area*

What if the sequence of Riemann sums we construct above converged to the same number I no matter what sequence of partitions whose norm goes to zero and associated evaluation sets we chose? Then, we would have that the value of this limit is *independent* of the choices above. This is indeed what we mean by the **Riemann Integral** of f on the interval $[a, b]$.

Definition 2.1.2. Riemann Integrability Of A Bounded Function

Let f be a bounded function on the finite interval $[a, b]$. if there is a number I so that

$$\lim_{n \to \infty} S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) = I$$

no matter what sequence of partitions $\{\boldsymbol{\pi_n}\}$ with associated sequence of evaluation sets $\{\boldsymbol{\sigma_n}\}$ we choose as long as $\lim_{n \to \infty} \| \boldsymbol{\pi_n} \| = 0$, we will say that the Riemann Integral of f on $[a, b]$ exists and equals the value I.

The value I is dependent on the choice of f and interval $[a, b]$. So we often denote this value by $I(f, [a, b])$ or more simply as, $I(f, a, b)$. Historically, the idea of the Riemann integral was developed using area approximation as an application, so the summing nature of the Riemann Sum was denoted by the 16^{th} century *letter* S which resembled an elongated or stretched letter S which looked like what we call the integral sign $\int$. Hence, the common notation for the Riemann Integral of f on $[a, b]$, when this value exists, is $\int_a^b f$. We usually want to remember what the independent variable of f is also and we want to remind ourselves that this value is obtained as we let the norm of the partitions go to zero. The symbol dx for the independent variable x is used as a reminder that $x_{i+1} - x_i$ is going to zero as the norm of the partitions goes to zero. So it has been very convenient to add to the symbol $\int_a^b f$ this information and use the augmented symbol $\int_a^b f(x)\, dx$ instead. Hence, if the independent variable was t instead of

x, we would use $\int_a^b f(t)\,dt$. Since for a function f, the name we give to the independent variable is a matter of personal choice, we see that the choice of variable name we use in the symbol $\int_a^b f(t)\,dt$ is very arbitrary. Hence, it is common to refer to the independent variable we use in the symbol $\int_a^b f(t)\,dt$ as the dummy variable of integration.

We need a few more facts. We shall prove later the following things are true about the Riemann Integral of a bounded function. First, we know when a bounded function actually has a Riemann integral from Theorem 2.1.1.

Theorem 2.1.1. Existence Of The Riemann Integral

Let f be a bounded function on the finite interval $[a,b]$. Then the Riemann integral of f on $[a,b]$, $\int_a^b f(t)dt$ exists if

1. *f is continuous on $[a,b]$*
2. *f is continuous except at a finite number of points on $[a,b]$.*

Further, if f and g are both Riemann integrable on $[a,b]$ and they match at all but a finite number of points, then their Riemann integrals match; i.e. $\int_a^b f(t)dt$ equals $\int_a^b g(t)dt$.

The function given by Equation 2.1 is bounded but continuous nowhere on $[-1,1]$ and it is indeed possible to prove it does not have a Riemann integral on that interval.

$$f(t) = \begin{cases} 1 & \text{if } t \text{ is a rational number} \\ -1 & \text{if } t \text{ is an irrational number} \end{cases} \tag{2.1}$$

However, most of the functions we want to work with do have a lot of *smoothness*, i.e. continuity and even differentiability on the intervals we are interested in. Hence, Theorem 2.1.1 will apply. Here are some examples:

1. If $f(t)$ is t^2 on the interval $[-2,4]$, then $\int_{-2}^4 f(t)dt$ does exist as f is continuous on this interval.
2. If g was defined by

$$g(t) = \begin{cases} t^2 & -2 \le t < 1 \text{ and } 1 < t \le 4 \\ 5 & t = 1 \end{cases}$$

   we see g is not continuous at only one point and so it is Riemann integrable on $[-2,4]$. Moreover, since f and g are both integrable and match at all but one point, their Riemann integrals are equal.

However, with that said, in this course, we want to relax the smoothness requirements on the functions f we work with and define a more general type of integral for this less restricted class of functions.

2.1.3 The Fundamental Theorem Of Calculus

There is a big connection between the idea of the *antiderivative* of a function f and its Riemann integral. For a positive function f on the finite interval $[a,b]$, we can construct the area under the curve function

$F(x) = \int_a^x f(t)\,dt$ where for convenience we choose an x in the open interval (a, b). We show $F(x)$ and $F(x + h)$ for a small positive h in Figure 2.4. Let's look at the difference in these areas:

$$\begin{aligned} F(x+h) - F(x) &= \int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt \\ &= \int_a^x f(t)\,dt + \int_x^{x+h} f(t)\,dt - \int_a^x f(t)\,dt \\ &= \int_x^{x+h} f(t)\,dt \end{aligned}$$

where we have used standard properties of the Riemann integral to write the first integral as two pieces and then do a subtraction. Now divide this difference by the change in x which is h. We find

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t)\,dt \tag{2.2}$$

The difference in area, $\int_x^{x+h} f(t)\,dt$, is the second shaded area in Figure 2.4. Clearly, we have

$$F(x+h) - F(x) = \int_x^{x+h} f(t)\,dt \tag{2.3}$$

We know that f is bounded on $[a, b]$; hence, there is a number B so that $f(t) \leq B$ for all t in $[a, b]$. Thus, using Equation 2.3, we see

$$F(x+h) - F(x) \leq \int_x^{x+h} B\,dt = B\,h \tag{2.4}$$

From this we can see that

$$\begin{aligned} \lim_{h \to 0} \left(F(x+h) - F(x)\right) &\leq \lim_{h \to 0} B\,h \\ &= 0 \end{aligned}$$

We conclude that F is continuous at each x in $[a, b]$ as

$$\lim_{h \to 0} \left(F(x+h) - F(x)\right) = 0$$

It seems that the new function F we construct by integrating the function f in this manner, always builds a new function that is continuous. Is F differentiable at x? If f is continuous at x, then given a positive ϵ, there is a positive δ so that

$$f(x) - \epsilon < f(t) < f(x) + \epsilon \text{ if } x - \delta < t < x + \delta$$

and t is in $[a, b]$. So, if h is less than δ, we have

$$\frac{1}{h} \int_x^{x+h} (f(x) - \epsilon) \; < \; \frac{F(x\,+\,h)\;-\;F(x)}{h} \;\; = \;\; \frac{1}{h} \int_x^{x+h} f(t)\,dt \; < \; \frac{1}{h} \int_x^{x+h} (f(x) + \epsilon)$$

This is easily evaluated to give

$$f(x) - \epsilon \; < \; \frac{F(x\,+\,h)\;-\;F(x)}{h} \;\; = \;\; \int_x^{x+h} f(t)\,dt \; < \; f(x) + \epsilon$$

if h is less than δ. This shows that

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

You should be able to believe that a similar argument would work for negative values of h: i.e.,

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

This tells us that $F'(x)$ exists and equals $f(x)$ as long as f is continuous at x as

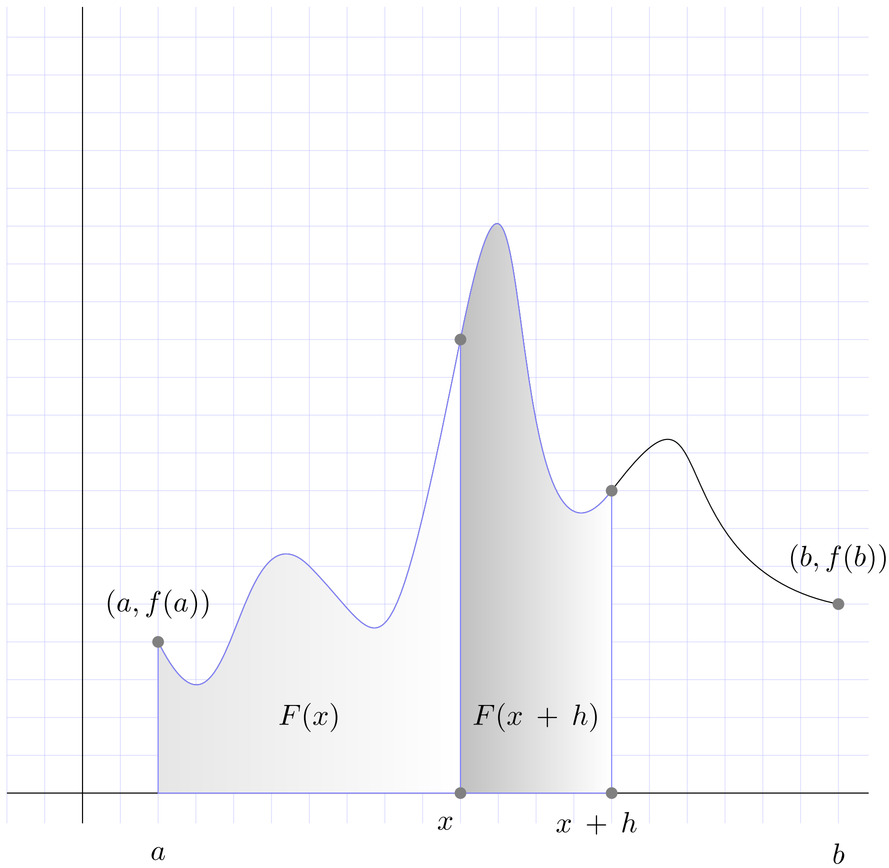
$$\begin{aligned} F'(x^+) \;\; &= \;\; \lim_{h \to 0^+} \frac{F(x\,+\,h)\;-\;F(x)}{h} \; = \; f(x) \\ F'(x^-) \;\; &= \;\; \lim_{h \to 0^-} \frac{F(x\,+\,h)\;-\;F(x)}{h} \; = \; f(x) \end{aligned}$$

This relationship is called **The Fundamental Theorem of Calculus**. The same sort of argument works for x equals a or b but we only need to look at the derivative from one side. We will prove this sort of theorem using fairly relaxed assumptions on f for the interval $[a, b]$ in the later Chapters. Even if we just consider the world of Riemann Integration, we only need to assume that f is Riemann Integrable on $[a, b]$ which allows for jumps in the function.

Theorem 2.1.2. Fundamental Theorem Of Calculus

Let f be Riemann Integrable on $[a, b]$. Then the function F defined on $[a, b]$ by $F(x) = \int_a^x f(t)dt$ satisfies

1. *F is continuous on all of $[a, b]$*
2. *F is differentiable at each point x in $[a, b]$ where f is continuous and $F'(x) = f(x)$.*



A generic curve f on the interval $[a, b]$ which is always positive. We let $F(x)$ be the area under this curve from a to x. This is indicated by the shaded region.

Figure 2.4: *The Function* $F(x)$

2.1.4 The Cauchy Fundamental Theorem Of Calculus

We can use the Fundamental Theorem of Calculus to learn how to evaluate many Riemann integrals. Let G be an antiderivative of the function f on $[a, b]$. Then, by definition, $G'(x) = f(x)$ and so we know G is continuous at each x. But we still don't know that f itself is continuous. However, if we assume f *is* continuous, then if we define F on $[a, b]$ by

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

the Fundamental Theorem of Calculus, Theorem 2.1.2, is applicable. Thus, $F'(x) = f(x)$ at each point. But that means $F' = G' = f$ at each point. Functions whose derivatives are the same must differ by a constant. Call this constant C. We thus have $F(x) = G(x) + C$. So, we have

$$\begin{aligned} F(b) &= f(a) + \int_a^b f(t)dt = G(b) + C \\ F(a) &= f(a) + \int_a^a f(t)dt = G(a) + C \end{aligned}$$

But $\int_a^a f(t)\,dt$ is zero, so we conclude after some rewriting

$$\begin{aligned} G(b) &= f(a) + \int_a^b f(t)dt + C \\ G(a) &= f(a) + C \end{aligned}$$

And after subtracting, we find the important result

$$G(b) - G(a) = \int_a^b f(t)dt$$

This is huge! This is what tells us how to integrate many functions. For example, if $f(t) = t^3$, we can guess the antiderivatives have the form $t^4/4 + C$ for an arbitrary constant C. Thus, since $f(t) = t^3$ is continuous, the result above applies. We can therefore calculate Riemann integrals like these:

1.
$$\begin{aligned} \int_1^3 t^3\,dt &= \left.\frac{t^4}{4}\right|_1^3 \\ &= \frac{3^4}{4} - \frac{1^4}{4} \\ &= \frac{80}{4} \end{aligned}$$

2.
$$\begin{aligned} \int_{-2}^4 t^3\,dt &= \left.\frac{t^4}{4}\right|_{-2}^4 \\ &= \frac{4^4}{4} - \frac{(-2)^4}{4} \\ &= \frac{256}{4} - \frac{16}{4} \\ &= \frac{240}{4} \end{aligned}$$

Let's formalize this as a theorem. All we really need to prove this result is that f is Riemann integrable on $[a, b]$, which is true if our function f is continuous.

Theorem 2.1.3. Cauchy Fundamental Theorem Of Calculus

Let G be any antiderivative of the Riemann integrable function f on the interval $[a, b]$. Then $G(b) - G(a) = \int_a^b f(t)\,dt$.

2.1.5 Applications

With the Cauchy Fundamental Theorem of Calculus under our belt, we can guess a lot of antiderivatives and from that know how to evaluate many Riemann integrals. Let's get started.

1. It is easy to guess the antiderivative of a power of t as we have already mentioned. We know the antiderivative of the following are easy to figure out:

   (a) If $f(t) = t^5$, then the antiderivative of f is any function of the form $F(t) = t^6/6 + C$ where C can be any constant.

   (b) If $f(t) = t^{-5}$, it is still easy to guess the antiderivative which is $F(t) = t^{-4}/(-4) + C$, where C is an arbitrary constant.

   The common symbol for the antiderivative of f has evolved to be $\int f$ because of the close connection between the antiderivative of f and the Riemann integral of f which is given in the Cauchy Fundamental Theorem of Calculus, Theorem 2.1.3. The usual Riemann integral, $\int_a^b f(t)\,dt$ of f on $[a,b]$ computes a *definite* value – hence, the symbol $\int_a^b f(t)\,dt$ is usually referred to as the *definite integral of f on $[a,b]$* to contrast it with the *family* of functions represented by the antiderivative $\int f$. Since the antiderivatives are arbitrary up to a constant, most of us refer to the antiderivative as the *indefinite integral of f* . Also, we hardly ever say "let's find the antiderivative of f" – instead, we just say, "let's *integrate* f". We will begin using this shorthand now! We can state these results as Theorem 2.1.4.

   Theorem 2.1.4. Antiderivatives Of Simple Powers

   If p is any power other than -1, then the antiderivative of $f(t) = t^p$ is $F(t) = t^{p+1}/(p+1)+C$. This is also expressed as $\int t^p\,dt = t^{p+1}/(p+1) + C$

2. The Riemann integral of the function f on $[a,b]$ can also be easily computed. We state this Theorem 2.1.5

   Theorem 2.1.5. Definite Integrals Of Simple Powers

   If p is any power other than -1, then the definite integral of $f(t) = t^p$ on $[a,b]$ is $\int_a^b t^p\,dt = t^{p+1}/(p+1)\Big|_a^b$

3. The simple trigonometric functions $\sin(t)$ and $\cos(t)$ also have straightforward antiderivatives as shown in Theorem 2.1.6.

   Theorem 2.1.6. Antiderivatives of Simple Trigonometric Functions

   (a) The antiderivative of $\sin(t)$ equals $-\cos(t) + C$

   (b) The antiderivative of $\cos(t)$ equals $\sin(t) + C$

4. The definite integrals of the sin and cos functions are then:

Theorem 2.1.7. Definite Integrals Of Simple Trigonometric Functions

(a) $\int_a^b \sin(t)\, dt$ *is* $-\cos(t)\Big|_a^b$

(b) $\int_a^b \cos(t)$ *is* $\sin(t)\Big|_a^b$

2.1.6 Simple Substitution Techniques

We can use the tools above to figure out how to *integrate* many functions that seem complicated but instead are just *disguised* versions of simple power function integrations. Let's go through some in great detail.

Exercise 2.1.1. *Compute* $\int (t^2 + 1)\, 2t\, dt$

Solution 2.1.1. *When you look at this integral, you should train yourself to see the simpler integral* $\int u\, du$ *where* $u(t) = t^2 + 1$. *Here are the steps:*

1. *We make the change of variable* $u(t) = t^2 + 1$. *Now differentiate both sides to see* $u'(t) = 2t$. *Thus, we have*

$$\int (t^2 + 1)\, 2t\, dt \quad = \quad \int u(t)\, u'(t)\, dt$$

2. *Now recall the chain rule for powers of functions, we know*

$$\left((u(t))^2\right)'(t) \quad = \quad 2\, u(t)\, u'(t)$$

Thus,

$$u(t)\, u'(t) \quad = \quad \frac{1}{2}\left((u(t))^2\right)'(t)$$

This then tells us that

$$\begin{aligned} \int (t^2 + 1)\, 2t\, dt \quad &= \quad \int u(t)\, u'(t)\, dt \\ &= \quad \int \frac{1}{2}\left((u(t))^2\right)'(t) dt \end{aligned}$$

Now, the notation $\int \left((u(t))^2\right)'(t) dt$ *is just our way of asking for the* antiderivative *of the function behind the integral sign. Here, that function is* $(u^2)'$. *This antiderivative is, of course, just* u^2*!*

Plugging that into the original problem, we find

$$\begin{aligned}\int (t^2 + 1)\, 2t\, dt &= \int u(t)\, u'(t)\, dt \\ &= \int \frac{1}{2} \left((u(t))^2\right)' (t) dt \\ &= \frac{1}{2}\, u^2(t) + C \\ &= \frac{1}{2}\, (t^2 + 1)^2 + C\end{aligned}$$

Whew!! That was awfully complicated looking. Let's do it again in a bit more streamlined fashion. Note all of the steps we go through below are the same as the longer version above, but since we write less detail down, it is much more compact. You need to get very good at understanding and doing all these steps!! Here is the second version:

Solution 2.1.2. 1. *We make the change of variable $u(t) = t^2 + 1$. But we write this more simply as $u = t^2 + 1$ so that the dependence of u on t is implied rather than explicitly stated. This simplifies our notation already! Now differentiate both sides to see $u'(t) = 2t$. We will write this as $du = 2t\, dt$, again hiding the t variable, using the fact that $\frac{du}{dt} = 2t$ can be written in its differential form (you should have seen this idea in your first Calculus course). Thus, we have*

$$\int (t^2 + 1)\, 2t\, dt = \int u\, du$$

2. *The antiderivative of u is $u^2/2 + C$ and so we have*

$$\begin{aligned}\int (t^2 + 1)\, 2t\, dt &= \int u\, du \\ &= \frac{1}{2}\, u^2 + C \\ &= \frac{1}{2}\, (t^2 + 1)^2 + C\end{aligned}$$

Now let's try one a bit harder:

Exercise 2.1.2. *Compute $\int (t^2 + 1)^3\, 4dt$*

Solution 2.1.3. *When you look at this integral, again you should train yourself to see the simpler integral $2 \int u^3\, du$ where $u(t) = t^2 + 1$. Here are the steps: first, the detailed version*

1. *We make the change of variable $u(t) = t^2 + 1$. Now differentiate both sides to see $u'(t) = 2t$. Thus, we have*

$$\int (t^2 + 1)^3\, 4dt = 2 \int u^3(t)\, u'(t)\, dt$$

2. *Now recall the chain rule for powers of functions, we know*

$$\left((u(t))^4\right)'(t) \quad = \quad 4\, u^3(t)\, u'(t)$$

Thus,

$$2\, u^3(t)\, u'(t) \quad = \quad 2\, \frac{1}{4}\left((u(t))^4\right)'(t)$$

This then tells us that

$$\begin{aligned} \int (t^2 + 1)^3\, 4dt &= 2 \int u^3(t)\, u'(t)\, dt \\ &= \int \frac{1}{2}\left((u(t))^4\right)'(t) dt \end{aligned}$$

Now, the notation $\int \left((u(t))^4\right)'(t)dt$ *is just our way of asking for the* antiderivative *of the function behind the integral sign. Here, that function is* $(u^4)'$. *This antiderivative is, of course, just* u^4*! Plugging that into the original problem, we find*

$$\begin{aligned} \int (t^2 + 1)^3\, 4dt &= 2 \int u^3(t)\, u'(t)\, dt \\ &= \frac{1}{2}\, u^4(t) + C \\ &= \frac{1}{2}\, (t^2 + 1)^4 + C \end{aligned}$$

Again, this was awfully complicated looking. the streamlined version is as follows:

1. *We make the change of variable* $u(t) = t^2 + 1$. *Now differentiate both sides to see* $u'(t) = 2t$ *and write this as* $du = 2t\, dt$. *Thus, we have*

$$\int (t^2 + 1)^3\, 4dt \quad = \quad 2 \int u^3\, du$$

2. *The antiderivative of* u^3 *is* $u^4/4 + C$ *and so we have*

$$\begin{aligned} \int (t^2 + 1)^3\, 4dt &= 2 \int u^3\, du \\ &= \frac{1}{2}\, u^4 + C \\ &= \frac{1}{2}\, (t^2 + 1)^4 + C \end{aligned}$$

Now let's do one the short way only.

Exercise 2.1.3. *Compute* $\int \sqrt{t^2 + 1}\, 3t\, dt$.

Solution 2.1.4. *When you look at this integral, again you should train yourself to see the simpler integral* $3/2 \int u^{1/2}\, du$ *where* $u(t) = t^2 + 1$. *Here are the steps: we know* $du = 2t\, dt$. *Thus*

$$\begin{aligned} \int \sqrt{t^2 + 1}\, 3t\, dt &= \frac{3}{2} \int u^{\frac{1}{2}}\, du \\ &= \frac{3}{2} \frac{1}{\frac{3}{2}} u^{\frac{3}{2}} + C \\ &= \frac{3}{2} \frac{2}{3} (t^2 + 1)^{\frac{3}{2}} + C \end{aligned}$$

Exercise 2.1.4. *Compute* $\int \sin(t^2 + 1)\, 5t\, dt$.

Solution 2.1.5. *When you look at this integral, again you should train yourself to see the simpler integral* $5/2 \int \sin(u)\, du$ *where* $u(t) = t^2 + 1$. *Here are the steps: we know* $du = 2t\, dt$. *Thus*

$$\begin{aligned} \int \sin(t^2 + 1)\, 5t\, dt &= \frac{5}{2} \int \sin(u)\, du \\ &= \frac{5}{2} (-\cos(u)) + C \\ &= -\frac{5}{2} \cos(t^2 + 1) + C \end{aligned}$$

Now let's do a definite integral:

Exercise 2.1.5. *Compute* $\int_1^5 (t^2 + 2t + 1)^2\, (t + 1)\, dt$.

Solution 2.1.6. *When you look at this integral, again you should train yourself to see the simpler integral* $1/2 \int u^2\, du$ *where* $u(t) = t^2 + 2t + 1$. *Here are the steps: we know* $du = (2t + 2) dt$. *Thus*

$$\int_1^5 (t^2 + 2t + 1)^2\, (t + 1)\, dt = \frac{1}{2} \int_{t=1}^{t=5} u^2\, du$$

where we label the bottom and top limit of the integral in terms of the t *variable to remind ourselves that the original integration was respect to* t. *Then,*

$$\begin{aligned} \frac{1}{2} \int_{t=1}^{t=5} u^2\, du &= \frac{1}{2} \frac{u^3}{3} \Big|_{t=1}^{t=5} \\ &= \frac{1}{2} \frac{1}{3} (t^2 + 1)^3 \Big|_1^5 \\ &= \frac{1}{6} \left((26)^3 - 2^3\right) \end{aligned}$$

We will prove general substitution theorems for Riemann Integrable functions later. But it is really just an application of the chain rule!

2.2 The Riemann Integral of Functions With Jumps

Now let's look at the Riemann integral of functions which have points of discontinuity.

2.2.1 Removable Discontinuity

Consider the function f defined on $[-2, 5]$ by

$$f(t) = \begin{cases} 2t & -2 \leq t < 0 \\ 1 & t = 0 \\ (1/5)t^2 & 0 < t \leq 5 \end{cases}$$

Let's calculate $F(t) = \int_{-2}^{t} f(s)\,ds$. This will have to be done in several parts because of the way f is defined.

1. On the interval $[-2, 0]$, note that f is continuous except at one point, $t = 0$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function $2t$ is continuous on this interval and so is also Riemann integrable. Then since f on $[-2, 0]$ and $2t$ match at all but one point on $[-2, 0]$, their Riemann integrals must match. Hence, if t is in $[-2, 0]$, we compute F as follows:

$$\begin{aligned} F(t) &= \int_{-2}^{t} f(s)\,ds \\ &= \int_{-2}^{t} 2s\,ds \\ &= s^2 \Big|_{-2}^{t} \\ &= t^2 - (-2)^2 = t^2 - 4 \end{aligned}$$

2. On the interval $[0, 5]$, note that f is continuous except at one point, $t = 0$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function $(1/5)t^2$ is continuous on this interval and is therefore also Riemann integrable. Then since f on $[0, 5]$ and $(1/5)t^2$ match at all but one point on $[0, 5]$, their Riemann integrals must match. Hence, if t is in $[0, 5]$, we compute F as follows:

$$\begin{aligned} F(t) &= \int_{-2}^{t} f(s)\,ds \\ &= \int_{-2}^{0} f(s)\,ds + \int_{0}^{t} f(s)\,ds \\ &= \int_{-2}^{0} 2s\,ds + \int_{0}^{t} (1/5)s^2\,ds \\ &= s^2 \Big|_{-2}^{0} + (1/15)s^3 \Big|_{0}^{t} \\ &= -4 + t^3/15 \end{aligned}$$

Thus, we have found that

$$F(t) = \begin{cases} t^2 - 4 & -2 \leq t < 0 \\ t^3/15 - 4 & 0 < t \leq 5 \end{cases}$$

Note, we didn't define F at $t = 0$ yet. Since f is Riemann Integrable on $[-2, 5]$, we know from the Fundamental Theorem of Calculus, Theorem 2.1.2, that F must be continuous. Let's check. F is clearly continuous on either side of 0 and we note that $\lim_{t \to 0^-} F(t)$ which is $F(0^-)$ is -4 which is exactly the value of $F(0^+)$. Hence, F is indeed continuous at 0 and we can write

$$F(t) = \begin{cases} t^2 - 4 & -2 \le t \le 0 \\ t^3/15 - 4 & 0 \le t \le 5 \end{cases}$$

What about the differentiability of F? The Fundamental Theorem of Calculus guarantees that F has a derivative at each point where f is continuous and at those points $F'(t) = f(t)$. Hence, we know this is true at all t except 0. Note at those t, we find

$$F'(t) = \begin{cases} 2t & -2 \le t < 0 \\ (1/5)t^2 & 0 < t \le 5 \end{cases}$$

which is exactly what we expect. Also, note $F'(0^-) = 0$ and $F'(0^+) = 0$ as well. Hence, since the right and left hand derivatives match, we see $F'(0)$ does exist and has the value 0. But this is not the same as $f(0) = 1$. Note, F is **not** the antiderivative of f on $[-2, 5]$ because of this mismatch.

2.2.2 Jump Discontinuity

Now consider the function f defined on $[-2, 5]$ by

$$f(t) = \begin{cases} 2t & -2 \le t < 0 \\ 1 & t = 0 \\ 2 + (1/5)t^2 & 0 < t \le 5 \end{cases}$$

Let's calculate $F(t) = \int_{-2}^{t} f(s)\, ds$. Again, this will have to be done in several parts because of the way f is defined.

1. On the interval $[-2, 0]$, note that f is continuous except at one point, $t = 0$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function $2t$ is continuous on this interval and hence is also Riemann integrable. Then since f on $[-2, 0]$ and $2t$ match at all but one point on $[-2, 0]$, their Riemann integrals must match. Hence, if t is in $[-2, 0]$, we compute F as follows:

$$\begin{aligned} F(t) &= \int_{-2}^{t} f(s)\, ds \\ &= \int_{-2}^{t} 2s\, ds \\ &= s^2 \Big|_{-2}^{t} \\ &= t^2 - (-2)^2 = t^2 - 4 \end{aligned}$$

2. On the interval $[0, 5]$, note that f is continuous except at one point, $t = 0$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function $2 + (1/5)t^2$ is continuous on this interval and so

is also Riemann integrable. Then since f on $[0,5]$ and $2 + (1/5)t^2$ match at all but one point on $[0,5]$, their Riemann integrals must match. Hence, if t is in $[0,5]$, we compute F as follows:

$$\begin{aligned} F(t) &= \int_{-2}^{t} f(s)\,ds \\ &= \int_{-2}^{0} f(s)\,ds + \int_{0}^{t} f(s)\,ds \\ &= \int_{-2}^{0} 2s\,ds + \int_{0}^{t} (2 + (1/5)s^2)\,ds \\ &= s^2\Big|_{-2}^{0} + (2s + (1/15)s^3)\Big|_{0}^{t} \\ &= -4 + 2t + t^3/15 \end{aligned}$$

Thus, we have found that

$$F(t) = \begin{cases} t^2 - 4 & -2 \leq t < 0 \\ -4 + 2t + t^3/15 & 0 < t \leq 5 \end{cases}$$

As before, we didn't define F at $t = 0$ yet. Since f is Riemann Integrable on $[-2,5]$, we know from the Fundamental Theorem of Calculus, Theorem 2.1.2, that F must be continuous. F is clearly continuous on either side of 0 and we note that $\lim_{t \to 0^-} F(t)$ which is $F(0^-)$ is -4 which is exactly the value of $F(0^+)$. Hence, F is indeed continuous at 0 and we can write

$$F(t) = \begin{cases} t^2 - 4 & -2 \leq t \leq 0 \\ -4 + 2t + t^3/15 & 0 \leq t \leq 5 \end{cases}$$

What about the differentiability of F? The Fundamental Theorem of Calculus guarantees that F has a derivative at each point where f is continuous and at those points $F'(t) = f(t)$. Hence, we know this is true at all t except 0. Note at those t, we find

$$F'(t) = \begin{cases} 2t & -2 \leq t < 0 \\ 2 + (1/5)t^2 & 0 < t \leq 5 \end{cases}$$

which is exactly what we expect. However, when we look at the one sided derivatives, we find $F'(0^-) = 0$ and $F'(0^+) = 2$. Hence, since the right and left hand derivatives do not match, we see $F'(0)$ does not exist. Finally, note F is **not** the antiderivative of f on $[-2,5]$ because of this mismatch.

2.2.3 Homework

Exercise 2.2.1. *Compute* $\int_{-3}^{t} f(s)\, ds$ *for*

$$f(t) = \begin{cases} 3t & -3 \le t < 0 \\ 6 & t = 0 \\ (1/6)t^2 & 0 < t \le 6 \end{cases}$$

1. *Graph* f *and* F *carefully labeling all interesting points.*
2. *Verify that* F *is continuous and differentiable at all points but* $F'(0)$ *does not match* $f(0)$ *and so* F *is not the antiderivative of* f *on* $[-3, 6]$

Exercise 2.2.2. *Compute* $\int_{0}^{t} f(s)\, ds$ *for*

$$f(t) = \begin{cases} -2t & 2 \le t < 5 \\ 12 & t = 5 \\ 3t - 25 & 5 < t \le 10 \end{cases}$$

1. *Graph* f *and* F *carefully labeling all interesting points.*
2. *Verify that* F *is continuous and differentiable at all points but* $F'(5)$ *does not match* $f(5)$ *and so* F *is not the antiderivative of* f *on* $[2, 10]$

Exercise 2.2.3. *Compute* $\int_{-3}^{t} f(s)\, ds$ *for*

$$f(t) = \begin{cases} 3t & -3 \le t < 0 \\ 6 & t = 0 \\ (1/6)t^2 + 2 & 0 < t \le 6 \end{cases}$$

1. *Graph* f *and* F *carefully labeling all interesting points.*
2. *Verify that* F *is continuous and differentiable at all points except* 0 *and so* F *is not the antiderivative of* f *on* $[-3, 6]$

Exercise 2.2.4. *Compute* $\int_{0}^{t} f(s)\, ds$ *for*

$$f(t) = \begin{cases} -2t & 2 \le t < 5 \\ 12 & t = 5 \\ 3t & 5 < t \le 10 \end{cases}$$

1. *Graph* f *and* F *carefully labeling all interesting points.*
2. *Verify that* F *is continuous and differentiable at all points except* 5 *and so* F *is not the antiderivative of* f *on* $[2, 10]$

Chapter 3

Functions Of Bounded Variation

Now that we have seen a quick overview of what Riemann Integration entails, let's go back and look at it very carefully. This will enable us to extend it to a more general form of integration called **Riemann - Stieljes**. From what we already know about Riemann integrals, the Riemann integral is a mapping ϕ which is linear and whose domain is some subspace of the vector space of all bounded functions. Let $B[a,b]$ denote this vector space which is a normed linear space using the usual infinity norm. The set of all Riemann Integrable Functions can be denoted by the symbol $RI[a,b]$ and we know it is a subspace of $B[a,b]$. We also know that the subspace $C[a,b]$ of all continuous functions on $[a,]$ is contained in $RI[a,b]$. In fact, if $PC[a,b]$ is the set of all functions on $[a,b]$ that are piecewise continuous, then $PC[a,b]$ is also a vector subspace contained in $RI[a,b]$. Hence, we know $\phi : RI[a,b] \subseteq B[a,b] \rightarrow \Re$ is a linear functional on the subspace $RI[a,b]$. Also, if f is not zero, then

$$\begin{aligned}
\frac{\mid \int_a^b f(t)\,dt \mid}{\| f \|_\infty} &\leq \frac{\int_a^b \mid f(t) \mid \,dt}{\| f \|_\infty} \\
&\leq \frac{\int_a^b \| f \|_\infty \,dt}{\| f \|_\infty} \\
&= b - a
\end{aligned}$$

Thus, we see that $\| \phi \|_{op}$ is finite and ϕ is a bounded linear functional on a subspace of $B[a,b]$ if we use the infinity norm on $RI[a,b]$. But of course, we can choose other norms. There are clearly many functions in $B[a,b]$ that do not fit nicely into the development process for the Riemann Integral. So let $NI[a,b]$ denote a new subspace of functions which contains $RI[a,b]$. We know that the Riemann integral satisfies an important idea in analysis called **limit interchange**. That is, if a sequence of functions $\{f_n\}$ from $RI[a,b]$ converges in infinity norm to f that the following facts hold:

1. f is also in $RI[a,b]$

2. the classic limit interchange holds:

$$\lim_{n \to \infty} \int_a^b f_n(t)\, dt \quad = \quad \int_a^b \left(\lim_{n \to \infty} f_n(t) \right) dt$$

We can say this more abstractly as this: if $f_n \to f$ in $|| \cdot ||_\infty$ in $RI[a,b]$, then f remains in $RI[a,b]$ and

$$\lim_{n \to \infty} \phi(f_n) \quad = \quad \phi\left(\lim_{n \to \infty} f_n \right)$$

But if we wanted to extend ϕ to the larger subspace $NI[a,b]$ in such a way that it remained a bounded linear functional, we would also want to know what kind of sequence convergence we should use in order for the interchange ideas to work. There are lots of questions:

1. Do we need to impose a norm on our larger subspace $NI[a,b]$?
2. Can we characterize the subspace $NI[a,b]$ in some fashion?
3. If the extension is called $\hat{\phi}$, we want to make sure that $\hat{\phi}$ is exactly ϕ when we restrict our attention to functions in $RI[a,b]$

Also, do we have to develop integration only on finite intervals $[a,b]$ of $\Re$? How do we even extend traditional Riemann integration to unbounded intervals of $\Re$? All of these questions will be answered in the upcoming chapters, but first we will see how far we can go with the traditional Riemann approach. We will also see where the Riemann integral approach breaks down and makes us start to think of more general tools so that we can get our work done.

3.1 Partitions

Definition 3.1.1. Partition

A **partition** *of the finite interval $[a,b]$ is a finite collection of points, $\{x_0, \ldots, x_n\}$, ordered so that $a = x_0 < x_1 < \cdots < x_n = b$. We denote the partition by $\boldsymbol{\pi}$ and call each point x_i a partition point.*

For each $j = 1, \ldots, n-1$, we let $\Delta x_j = x_{j+1} - x_j$. The collection of all finite partitions of $[a,b]$ is denoted $\Pi[a,b]$.

Definition 3.1.2. Partition Refinements

The partition $\pi_1 = \{y_0, \ldots, y_m\}$ is said to be a **refinement** *of the partition $\pi_2 = \{x_0, \ldots, x_n\}$ if every partition point $x_j \in \pi_2$ is also in π_1. If this is the case, then we write $\pi_2 \preceq \pi_1$, and we say that π_1 is finer than π_2 or π_2 is coarser than π_1.*

Definition 3.1.3. Common Refinement

Given $\pi_1, \pi_2 \in \Pi[a,b]$, there is a partition $\pi_3 \in \Pi[a,b]$ which is formed by taking the union of π_1 and π_2 and using common points only once. We call this partition the common refinement of π_1 and π_2 and denote it by $\pi_3 = \pi_1 \vee \pi_2$.

Comment 3.1.1. *The relation $\preceq$ is a partial ordering of $\Pi[a,b]$. It is not a total ordering, since not all partitions are comparable. There is a coarsest partition, also called the trivial partition. It is given by $\pi_0 = \{a, b\}$. We may also consider uniform partitions of order k. Let $h = (b-a)/k$. Then $\pi = \{x_0 = a, x_0 + h, x_0 + 2h, \ldots, x_{k-1} = x_0 + (k-1)h, x_k = b\}$.*

Proposition 3.1.1. Refinements and Common Refinements

If $\pi_1, \pi_2 \in \Pi[a,b]$, then $\pi_1 \preceq \pi_2$ if and only if $\pi_1 \vee \pi_2 = \pi_2$.

Proof. *If $\pi_1 \preceq \pi_2$, then $\pi_1 = \{x_0, \ldots, x_p\} \subset \{y_0, \ldots, y_q\} = \pi_2$. Thus, $\pi_1 \cup \pi_2 = \pi_2$, and we have $\pi_1 \vee \pi_2 = \pi_2$. Conversely, suppose $\pi_1 \vee \pi_2 = \pi_2$. By definition, every point of π_1 is also a point of $\pi_1 \vee \pi_2 = \pi_2$. So, $\pi_1 \preceq \pi_2$.* ■

Definition 3.1.4. The Gauge or Norm of a Partition

For $\pi \in \Pi[a,b]$, we define the **gauge** *of π, denoted $\|\pi\|$, by $\|\pi\| = \max\{\Delta x_j \,:\, 1 \leq j \leq p\}$.*

3.1.1 Homework

Exercise 3.1.1. *Prove that the relation $\preceq$ is a partial ordering of $\Pi[a,b]$.*

Exercise 3.1.2. *Fix $\pi_1 \in \Pi[a,b]$. The set $C(\pi_1) = \{\pi \in \Pi[a,b] \,:\, \pi_1 \preceq \pi\}$ is called the* **core** *determined by π_1. It is the set of all partitions of $[a,b]$ that contain (or are finer than) π_1.*

1. *Prove that if $\pi_1 \preceq \pi_2$, then $C(\pi_2) \subset C(\pi_1)$.*
2. *Prove that if $\|\pi_1\| < \epsilon$, then $\|\pi\| < \epsilon$ for all $\pi \in C(\pi_1)$.*
3. *Prove that if $\|\pi_1\| < \epsilon$ and $\pi_2 \in \Pi[a,b]$, then $\|\pi_1 \vee \pi_2\| < \epsilon$.*

3.2 Monotone Functions

In our investigations of how monotone functions behave, we will need two fundamental facts about *infimum* and *supremum* of a set of numbers which are given in Lemma 3.2.1 and Lemma 3.2.2.

Lemma 3.2.1. The Infimum Tolerance Lemma

Let S be a nonempty set of numbers that is bounded below. Then given any tolerance ϵ, there is at least one element s in S so that

$$\inf(S) \;\leq\; s \;<\; \inf(S) \;+\; \epsilon$$

Proof. *This is an easy proof by contradiction. Assume there is some ϵ so that no matter what s from S we choose, we have*

$$s \;\geq\; \inf(S) \;+\; \epsilon$$

This says that $\inf(S) + \epsilon$ is a lower bound for S and so by definition, $\inf(S)$ must be bigger than or equal to this lower bound. But this is clearly not possible. So the assumption that such a tolerance ϵ exists is wrong and the conclusion follows. ■

and

Lemma 3.2.2. The Supremum Tolerance Lemma

Let T be a nonempty set of numbers that is bounded above. Then given any tolerance ϵ, there is at least one element t in T so that

$$\sup(T) - \epsilon < t \leq \sup(T)$$

Proof. *This again is an easy proof by contradiction and we include it for completeness. Assume there is some ϵ so that no matter what t from T we choose, we have*

$$t \leq \sup(T) - \epsilon$$

This says that $\sup(T) - \epsilon$ is an upper bound for T and so by definition, $\sup(T)$ must be less than or equal to this upper bound. But this is clearly not possible. So the assumption that such a tolerance ϵ exists is wrong and the conclusion must follow. ■

We are now in a position to discuss carefully monotone functions and other functions built from them. We follow discussions in (Douglas (2) 1996) at various places.

Definition 3.2.1. Monotone Functions

A real-valued function $f : [a, b] \to \mathbb{R}$ is said to be increasing (respectively, strictly increasing) if $x_1, x_2 \in [a, b]$, $x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$ (respectively, $f(x_1) < f(x_2)$). Similar definitions hold for decreasing and strictly decreasing functions.

Theorem 3.2.3. A Monotone Function Estimate

Let f be increasing on $[a, b]$, and let $\pi = \{x_0, \ldots, x_p\}$ be in $\Pi[a, b]$. For any $c \in [a, b]$, define

$$f(c^+) = \lim_{x \to c^+} f(x) \qquad \text{and} \qquad f(c^-) = \lim_{x \to c^-} f(x),$$

where we define $f(a^-) = f(a)$ and $f(b^+) = f(b)$. Then

$$\sum_{j=0}^{p} [f(x_j^+) - f(x_j^-)] \leq f(b) - f(a).$$

Proof. *First, we note that $f(x^+)$ and $f(x^-)$ always exist. The proof of this is straightforward. For $x \in (a, b]$, let $T_x = \{f(y) \, : \, a \leq y < x\}$. Then T_x is bounded above by $f(x)$, since f is monotone increasing. Hence, T_x has a well-defined supremum. Let $\epsilon > 0$ be given. Then, using the Supremum Tolerance Lemma, Lemma 3.2.2, there is a $y^* \in [a, x)$ such that $\sup T_x - \epsilon < f(y^*) \leq \sup T_x$. For any $y \in (y^*, x)$, we have $f(y^*) \leq f(y)$ since f is increasing. Thus, $0 \leq (\sup T_x - f(y)) \leq (\sup T_x - f(y^*)) < \epsilon$ for $y \in (y^*, x)$. Let $\delta = (x - y^*)/2$. Then, if $0 < x - y < \delta$, $\sup T_x - f(y) < \epsilon$. Since ϵ was arbitrary, this shows that $\lim_{y \to x^-} f(y) = supT_x$. The proof for $f(x^+)$ is similar, using the Infimum Tolerance Lemma, Lemma 3.2.1. You should be able to see that $f(x^-)$ is less than or equal to $f(x^+)$ for all x. We will define $f(a^-) = f(a)$ and $f(b^+) = f(b)$ since f is not defined prior to a or after b.*

To prove the stated result holds, first choose an arbitrary $y_j \in (x_j, x_{j+1})$ *for each* $j = 0, \ldots, p-1$. *Then, since* f *is increasing, for each* $j = 1, \ldots, p$, *we have* $f(y_{j-1}) \leq f(x_j^-) \leq f(x_j^+) \leq f(y_j)$. *Thus,*

$$f(x_j^+) - f(x_j^-) \quad \leq \quad f(y_j) - f(y_{j-1}). \tag{3.1}$$

We also have $f(a) \leq f(a^+) \leq f(y_0)$ *and* $f(y_{p-1}) \leq f(b^-) \leq f(b)$. *Thus, it follows that*

$$\begin{aligned}
\sum_{j=0}^{p} \big(f(x_j^+) - f(x_j^-)\big) \quad &= \quad f(x_0^+) - f(x_0^-) + \sum_{j=1}^{p-1} [f(x_j^+ - f(x_j^-)] \ + \ f(x_p^+) - f(x_p^-) \\
&\leq \quad f(a^+) - f(a^-) + \sum_{j=1}^{p-1} [f(y_j - f(y_{j-1})] + f(b^+) - f(b^-)
\end{aligned}$$

using Equation 3.1 and replacing x_0 *by* a *and* x_p *with* b. *We then note the sum on the right hand side collapses to* $f(y_{p-1}) - f(y_0)$. *Finally, since* $f(a^-) = f(a)$ *and* $f(b^+) = f(b)$, *we obtain*

$$\begin{aligned}
\sum_{j=0}^{p} \big(f(x_j^+) - f(x_j^-)\big) \quad &\leq \quad f(a^+) - f(a) + f(y_{p-1}) - f(y_0) + f(b) - f(b^-) \\
&\leq \quad f(y_0) - f(a) + f(y_{p-1}) - f(b^-) + f(b) - f(y_0) \\
&\leq \quad f(b) - f(a) + f(y_{p-1}) - f(b^-).
\end{aligned}$$

But $f(y_{p-1}) - f(b^-) \leq 0$, *so*

$$\sum_{j=0}^{p} \big(f(x_j^+) - f(x_j^-)\big) \quad \leq \quad f(b) - f(a).$$

■

Theorem 3.2.4. A Monotone Function Has A Countable Number of Discontinuities

If f *is monotone on* $[a, b]$, *the set of discontinuities of* f *is countable.*

Proof. *For concreteness, we assume* f *is monotone increasing. The decreasing case is shown similarly. Since* f *is monotone increasing, the only types of discontinuities it can have are jump discontinuities. If* $x \in [a, b]$ *is a point of discontinuity, then the size of the jump is given by* $f(x^+) - f(x^-)$. *Define* $D_k = \{x \in (a, b) \ : \ f(x^+) - f(x^-) > 1/k\}$, *for each integer* $k \geq 1$. *We want to show that* D_k *is finite.*

Select any finite subset S *of* D_k *and label the points in* S *by* $x_1, \ldots, x_p$ *with* $x_1 < x_2 < \cdots < x_p$. *If we add the point* $x_0 = a$ *and* $x_{p+1} = b$, *these points determine a partition* π. *Hence, by Theorem 3.2.3, we know that*

$$\sum_{j=1}^{p} [f(x_j^+) - f(x_j^-)] \quad \leq \quad \sum_{\pi} [f(x_j^+) - f(x_j^-)] \ \leq \ f(b) - f(a).$$

But each jump satisfies $f(x_j^+) - f(x_j^-) > 1/k$ and there are a total of p such points in S. Thus, we must have

$$p/k \;<\; \sum_{j=1}^{p} [f(x_j^+) - f(x_j^-)] \;\leq\; f(b) - f(a).$$

Hence, $p/k < f(b) - f(a)$, implying that $p < k[f(b) - f(a)]$. Thus, the cardinality of S is bounded above by the fixed constant $k[f(b) - f(a)]$. Let $\hat{N}$ be the first positive integer bigger than or equal to $k[f(b) - f(a)]$. If the cardinality of D_k were infinite, then there would be a subset T of D_k with cardinality $\hat{N} + 1$. The argument above would then tell us that $\hat{N} + 1 \leq k[f(b) - f(a)] \leq \hat{N}$ giving a contradiction. Thus, D_k must be a finite set. This means that $D = \cup_{k=1}^{\infty} D_k$ is countable also.

Finally, if x is a point where f is not continuous, then $f(x^+) - f(x^-) > 0$. Hence, there is a positive integer k_0 so that $f(x^+) - f(x^-) > 1/k_0$. This means x is in D_{k_0} and so is in D. ■

Definition 3.2.2. The Discontinuity Set Of A Monotone Function

Let f be monotone increasing on $[a, b]$. We will let S denote the set of discontinuities of f on $[a, b]$. We know this set is countable by Theorem 3.2.4 so we can label it as $S = \{x_j\}$. Define functions u and v on $[a, b]$ by

$$u(x) = \begin{cases} 0, & x = a \\ f(x) - f(x^-), & x \in (a, b] \end{cases}$$

$$v(x) = \begin{cases} f(x^+) - f(x), & x \in [a, b) \\ 0, & x = b \end{cases}$$

In Figure 3.1, we show a monotone increasing function with several jumps. You should be able to compute u and v easily at these jumps.

There are several very important points to make about these functions u and v which are listed below.

Comment 3.2.1.

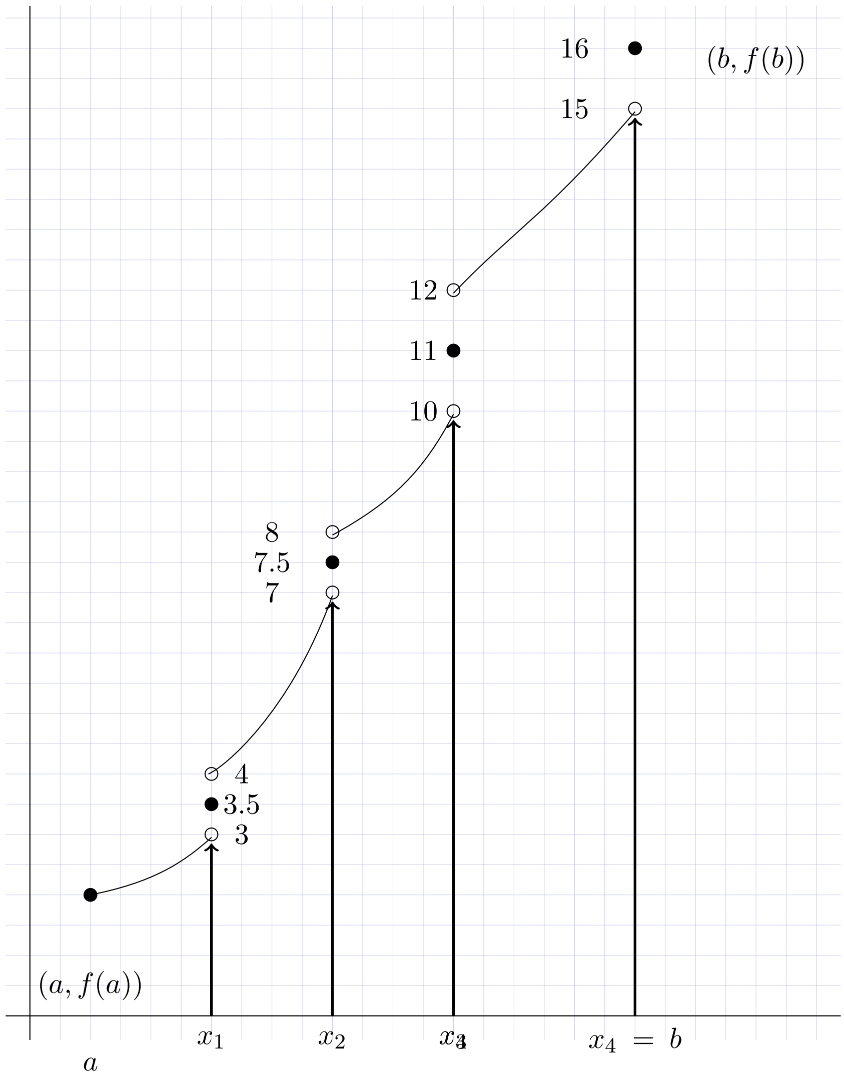
1. *Note that $u(x)$ is the left-hand jump of f at $x \in (a, b]$ and $v(x)$ is the right-hand jump of f at $x \in [a, b)$.*

2. *Both u and v are nonnegative functions and $u(x) + v(x) = f(x^+) - f(x^-)$ is the total jump in f at x, for $x \in (a, b)$.*

3. *Moreover, f is continuous at x from the left if and only if $u(x) = 0$, and f is continuous from the right at x if and only if $v(x) = 0$.*

4. *Finally, f is continuous on $[a, b]$ if and only if $u(x) = v(x) = 0$ on $[a, b]$.*

Now, let S_0 be any finite subset of S. From Theorem 3.2.3, we have



A generic curve f on the interval $[a, b]$ which is always positive. We show four points of discontinuity x_1, x_2, x_3 and x_4. Note $u(x_1) = 0.5$, $u(x_2) = 0.5$, $u(x_3) = 1$ and $u(x_4) = 1$. Also, we see $v(x_1) = 0.5$, $v(x_2) = 0.5$, $v(x_3) = 1$ and $v(x_4) = 0$.

Figure 3.1: *The Function $F(x)$*

$$\sum_{x \in S_0} f(x^+) - f(x^-) \quad \leq \quad f(b) - f(a)$$

This implies

$$\begin{aligned} \sum_{x \in S_0} u(x) + v(x) &\leq f(b) - f(a) \\ \sum_{x \in S_0} u(x) + \sum_{x \in S_0} v(x) &\leq f(b) - f(a). \end{aligned}$$

The above tells us that the set of numbers we get by evaluating this sum over finite subsets of S is bounded above by the number $f(b) - f(a)$. Hence, $\sum_{j=1}^{n} u(x_j)$ and $\sum_{j=1}^{n} v(x_j)$ are bounded above by $f(b) - f(a)$ for all n. Thus, these sets of numbers have a finite supremum. But u and v are nonnegative functions, so these sums form monotonically increasing sequences. Hence, these sequences converge to their supremum which we label as $\sum_{j=1}^{\infty} u(x_j)$ and $\sum_{j=1}^{\infty} v(x_j)$.

Now, consider a nonempty subset, T, of $[a, b]$, and suppose $F \subset S \cap T$ is finite. Then, by the arguments already presented, we know that

$$\sum_{x_j \in F} u(x_j) + \sum_{x_j \in F} v(x_j) \leq f(b) - f(a). \tag{3.2}$$

This implies

$$\sum_{x_j \in F} u(x_j) \;\leq\; f(b) - f(a) \qquad \text{and} \qquad \sum_{x_j \in F} v(x_j) \;\leq\; f(b) - f(a).$$

From this, it follows that

$$\sum_{x_j \in S \cap T} u(x_j) \quad = \quad \sup\{ \sum_{x_j \in F} u(x_j) \; : \; F \subset S \cap T, \; F \text{ finite} \}.$$

Likewise, we also have

$$\sum_{x_j \in S \cap T} v(x_j) \quad = \quad \sup\{ \sum_{x_j \in F} v(x_j) \; : \; F \subset S \cap T, F \text{ finite} \}.$$

Definition 3.2.3. The Saltus Function Associated With A Monotone Function
For $x, y \in [a, b]$ with $x < y$, define

$$S[x, y] \;=\; S \cap [x, y], \;\; S[x, y) \;=\; S \cap [x, y), \;\; S(x, y] \;=\; S \cap (x, y] \;\; \textit{and} \;\; S(x, y) \;=\; S \cap (x, y)$$

Then, define the function $S_f : [a, b] \rightarrow \mathbb{R}$ by

$$S_f(x) = \begin{cases} f(a), & x = a \\ f(a) + \sum_{x_j \in S(a,x]} u(x_j) \; + \; \sum_{x_j \in S[a,x)} v(x_j), & a \; < \; x \leq b \end{cases}$$

We call S_f the **Saltus Function** *associated with the monotone increasing function f.*

Intuitively, $S_f(x)$ is the sum of all of the jumps (i.e. discontinuities) up to and including the left-hand jump at x. In essence, it is a generalization of the idea of a *step function*.

Theorem 3.2.5. Properties of The Saltus Function

Let $f : [a, b] \to \mathbb{R}$ be monotone increasing. Then

1. *S_f is monotone increasing on $[a, b]$;*
2. *if $x < y$, with $x, y \in [a, b]$, then $0 \leq S_f(y) - S_f(x) \leq f(y) - f(x)$;*
3. *S_f is continuous on $S^c \cap [a, b]$ where S^c is the complement of the set S.*

Proof. *Suppose $x < y$. Then*

$$\begin{aligned} S_f(y) - S_f(x) &= \sum_{x_j \in S(a,y]} u(x_j) - \sum_{x_j \in S(a,x]} u(x_j) + \sum_{x_j \in S[a,y)} v(x_j) - \sum_{x_j \in S[a,x)} v(x_j) \\ &= \sum_{x_j \in S(x,y]} u(x_j) + \sum_{x_j \in S[x,y)} v(x_j) \\ &\geq 0. \end{aligned}$$

This proves the first statement. Now, suppose $x, y \in [a, b]$ with $x < y$. Let F be a subset of $[a, b]$ that consists of a finite number of points of the form $F = \{x_0 = x, x_1, \ldots, x_p = y\}$, such that $x = x_0 < x_1 < \cdots < x_p = y$. In other words, F is a partition of $[x, y]$. Then, by Equation 3.2 we know

$$\sum_{x_j \in F \cap S(x,y]} u(x_j) + \sum_{x_j \in F \cap S[x,y)} v(x_j) \leq f(y) - f(x).$$

Taking the supremum of the left-hand side over all such sets, F, we obtain

$$\sum_{x_j \in S(x,y]} u(x_j) + \sum_{x_j \in S[x,y)} v(x_j) \leq f(y) - f(x).$$

But by the remarks made in the first part of this proof, this sum is exactly $S_f(y) - S_f(x)$. We conclude that $S_f(y) - S_f(x) \leq f(y) - f(x)$ as desired.

Finally, let x be a point in $S^c \cap [a, b]$. Then f is continuous at x, so, given $\epsilon > 0$, there is a $\delta > 0$ such that $y \in [a, b]$ and $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$. But by the second part of this proof, we have $|S_f(x) - S_f(y)| \leq |f(y) - f(x)| < \epsilon$. Thus, S_f is continuous at x. ∎

So, why do we care about S_f? The function S_f measures, in a sense, the degree to which f fails to be continuous. If we subtract S_f from f, we would be subtracting its discontinuities, resulting in a continuous function that behaves similarly to f.

Definition 3.2.4. The Continuous Part of A Monotone Function

Define $f_c : [a, b] \to \mathbb{R}$ by $f_c(x) = f(x) - S_f(x)$.

Theorem 3.2.6. Properties of f_c

1. *f_c is monotone on $[a, b]$.*
2. *f_c is continuous also.*

Proof. *The proof that f_c is monotone is left to you as an exercise with this generous hint:*

Hint. *Note if $x < y$ in $[a, b]$, then*

$$f_c(y) - f_c(x) = (f(y) - f(x)) - (S_f(y) - S_f(x)).$$

The right hand side is non negative by Theorem 3.2.5. □

To prove f_c is continuous is a bit tricky. We will do most of the proof but leave a few parts for you to fill in.

Pick any x in $[a, b)$ and any positive ϵ. Since the $f(x^+)$ exists, there is a positive δ so that $0 \leq f(y) - f(x^+) < \epsilon$ if $x < y < x + \delta$. Thus, for such y,

$$\begin{aligned} f_c(y) - f_c(x) &= [f(y) - S_f(y)] - [f(x) - S_f(x)] \\ &= f(y) - \left\{ \sum_{x_j \in S(a,y]} u(x_j) + \sum_{x_j \in S[a,y)} v(x_j) \right\} \\ &\quad - f(x) + \left\{ \sum_{x_j \in S(a,x]} u(x_j) + \sum_{x_j \in S[a,x)} v(x_j). \right\} \end{aligned}$$

Recall, $S(a, y] = S(a, x] \cup S(x, y]$ and $S[a, y) = S[a, x) \cup S[x, y)$. So,

$$f_c(y) - f_c(x) = f(y) - \left\{ \sum_{x_j \in S(x,y]} u(x_j) + \sum_{x_j \in S[x,y)} v(x_j) \right\} - f(x)$$

Now, the argument reduces to two cases:

1. *if y and x are points of discontinuity, we get*

$$\begin{aligned} f_c(y) - f_c(x) &= f(y) - u(y) - \left\{ \sum_{x_j \in S(x,y)} u(x_j) + \sum_{x_j \in S(x,y)} v(x_j) \right\} - f(x) - v(x) \\ &= f(y) - (f(y) - f(y^-)) - \left\{ \sum_{x_j \in S(x,y)} u(x_j) + \sum_{x_j \in S(x,y)} v(x_j) \right\} - f(x) - (f(x^+) - f(x)) \\ &\leq f(y^-) - f(x^+) \\ &\leq f(y) - f(x^+) < \epsilon \end{aligned}$$

2. *if either x and/ or y are not a point of discontinuity, a similar argument holds*

Thus, we see f_c is continuous from the right at this x. Now use a similar argument to show continuity from the left at x. Together, these arguments show f_c is continuous at x. ■

3.2.1 Worked Out Example

Let's define f on $[0,2]$ by

$$f(x) = \begin{cases} -2 & x = 0 \\ x^3 & 0 < x < 1 \\ 9/8 & x = 1 \\ x^4/4 + 1 & 1 < x < 2 \\ 7 & x = 2 \end{cases}$$

1. Find u and v

2. Find S_f

3. Find f_c

4. Following the discussion in Section 2.2 explain how to compute the Riemann Integral of f and find its value (yes, this is in the careful rigorous section and so this problem is a bit out of place, but we will be dotting all of our i's and crossing all of our t's soon enough!)

Solution 3.2.1. *First, note $f(0^-) = -2$, $f(0) = -2$ and $f(0^+) = 0$ and so 0 is a point of discontinuity. Further, $f(1^-) = 1$, $f(1) = 2$ and $f(1^+) = 5/4$ giving another point of discontinuity at 1. Finally, since $f(2^-) = 5$, $f(2) = 7$ and $f(2^+) = 7$, there is a third point of discontinuity at 2. So, the set of discontinuities of f is $S = \{0,1,2\}$. Thus,*

$$S(0,x] = \begin{cases} \emptyset & 0 < x < 1 \\ \{1\} & 1 \leq x < 2 \\ \{1,2\} & 2 = x \end{cases} \quad \textit{and } S[0,x) = \begin{cases} \{0\} & 0 < x \leq 1 \\ \{0,1\} & 1 < x \leq 2 \end{cases}$$

Also,

$$u(x) = \begin{cases} 0 & x = 0 \\ 0 & 0 < x < 1 \\ 9/8 - 1 = 1/8 & x = 1 \\ 0 & 1 < x < 2 \\ 7 - 5 = 2 & 2 = x \end{cases} \quad \textit{and } v(x) = \begin{cases} 0 - (-2) = 2 & x = 0 \\ 0 & 0 < x < 1 \\ 5/4 - 9/8 = 1/8 & x = 1 \\ 0 & 1 < x < 2 \\ 0 & 2 = x \end{cases}$$

Now, here

$$S_f(x) = \begin{cases} f(0) = -2, & x = 0 \\ f(0) + \sum_{x_j \in S(0,x]} u(x_j) + \sum_{x_j \in S[0,x)} v(x_j) & 0 < x \leq 2 \end{cases}$$

Thus,

$$S_f(x) = \begin{cases} -2, & x = 0 \\ -2 + v(0) = -2 + 2 = 0 & 0 < x < 1 \\ -2 + u(1) + v(0) = -2 + 1/8 + 2 = 1/8 & x = 1 \\ -2 + u(1) + v(0) + v(1) = -2 + 1/8 + 2 + 1/8 = 1/4 & 1 < x < 2 \\ -2 + u(1) + u(2) + v(0) + v(1) = -2 + 1/8 + 2 + 2 + 1/8 = 9/4 & x = 2 \end{cases}$$

So, S_f is the nice step function and $f_c = f - S_f$ gives

$$S_f(x) = \begin{cases} -2, & x = 0 \\ 0 & 0 < x < 1 \\ 1/8 & x = 1 \\ 1/4 & 1 < x < 2 \\ 9/4 & x = 2 \end{cases} \quad \textit{and } f_c(x) = \begin{cases} -2 - (-2) = 0 & x = 0 \\ x^3 - 0 = x/3 & 0 < x < 1 \\ 9/8 - 1/8 = 1 & x = 1 \\ x^4/4 + 1 - 1/4 = x^4/4 + 3/4 & 1 < x < 2 \\ 7 - 9/4 = 19/4 & x = 2 \end{cases}$$

We see f_c is continuous on $[0, 2]$. Finally, we can compute the Riemann integral of f on $[0, 2]$.

Let's calculate $F(t) = \int_0^t f(x)\,dx$. This will have to be done in several parts because of the way f is defined.

1. *On the interval $[0, 1]$, note that f is continuous except at two points, $x = 0$ and $x = 1$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function x^3 is continuous on this interval and so is also Riemann integrable. Then since f on $[0, 1]$ and x^3 match at all but two points on $[0, 2]$, their Riemann integrals must match. Hence, if t is in $[-2, 0]$, we compute F as follows:*

$$\begin{aligned} F(t) &= \int_0^t f(x)\,dx \\ &= \int_0^t x^3\,dx \\ &= x^4/4 \Big|_0^t \\ &= t^4/4 \end{aligned}$$

2. *On the interval $[1, 2]$, note that f is continuous except at the two points, $x = 1$ and $x = 2$. Hence, f is Riemann integrable by Theorem 2.1.1. Also, the function $1 + x^4/4$ is continuous on this interval and so is also Riemann integrable. Then since f on $[1, 2]$ and $1 + x^4/4$ match at all but two points on $[1, 2]$, their Riemann integrals must match. Hence, if t is in $[1, 2]$, we compute*

F as follows:

$$\begin{aligned} F(t) &= \int_0^t f(x)\,dx \\ &= \int_0^1 f(x)\,dx + \int_1^t f(s)\,ds \\ &= \int_0^1 x^3\,dx + \int_1^t (1 + x^4/4)\,dx \\ &= x^4/4\,|_0^1 + (x + x^5/5)\,|_1^t \\ &= 1/4 + (t + t^5/5) - (1 + 1/5) \\ &= t^5/5 + t - 19/20 \end{aligned}$$

Thus, we have found that

$$F(t) = \begin{cases} t^4/4 & 0 \le t \le 1 \\ t^5/5 + t - 19/20 & 1 \le t \le 2 \end{cases}$$

Note, we know from the Fundamental Theorem of Calculus, Theorem 2.1.2, that F must be continuous. To check this at an interesting point such as $t = 1$, note F is clearly continuous on either side of 1 and we note that $\lim_{t \to 1^-} F(t)$ which is $F(1^-)$ is $1/4$ which is exactly the value of $F(1^+)$. Hence, F is indeed continuous at 1!

What about the differentiability of F? The Fundamental Theorem of Calculus guarantees that F has a derivative at each point where f is continuous and at those points $F'(t) = f(t)$. Hence, we know this is true at all t except 0, 1 and 2 because these are points of discontinuity of f. F' is nicely defined at 0 and 1 as a one sided derivative and at all other t save 1 by

$$F'(t) = \begin{cases} t^3 & 0 \le t < 1 \\ t^4 + 1 & 0 < t \le 2 \end{cases}$$

However, when we look at the one sided derivatives, we find $F'(0^+) = 0 \ne f(0) = -2$, $F'(2^-) = 17 \ne f(2) = 7$ and $F'(1^-) = 1$ and $F'(1^+) = 2$ giving $F'(1)$ does not even exist. Thus, note F is **not** *the antiderivative of f on $[0, 2]$ because of this mismatch.*

3.2.2 Homework

Exercise 3.2.1. *Prove f_c is monotone.*

Exercise 3.2.2. *Let's define f on $[0,2]$ by*

$$f(x) = \begin{cases} -1 & x = 0 \\ x^2 & 0 < x < 1 \\ 7/4 & x = 1 \\ \sqrt{x+3} & 1 < x < 2 \\ 3 & x = 2 \end{cases}$$

1. *Find u and v*
2. *Find S_f*
3. *Find f_c*
4. *Do a nice graph of u, v, f, f_c and S_f*
5. *Following the discussion in Section 2.2 explain how to compute the Riemann Integral of f and find its value (yes, this is in the careful rigorous section and so this problem is a bit out of place, but we will be dotting all of our i's and crossing all of our t's soon enough!)*

3.3 Functions of Bounded Variation

The next important topic for us is to consider the class of functions of *bounded variation*. We will develop this classically here, but in later chapters, we will define similar concepts using abstract measures. We are going to find out that functions of bounded variation can also be represented as the difference of two increasing functions and that there classical derivative exists everywhere except a set of measure zero (yes, that idea is not defined yet, but I believe in teasers!). Let's get on with it.

Definition 3.3.1. Functions Of Bounded Variation
Let $f : [a,b] \to \mathbb{R}$ and let $\pi \in \Pi[a,b]$ be given by $\pi = \{x_0 = a, x_1, \ldots, x_p = b\}$. Define $\Delta f_j = f(x_j) - f(x_{j-1})$ for $1 \leq j \leq p$. If there exists a positive real number, M, such that

$$\sum_{\pi} |\Delta f_j| \leq M$$

for all $\pi \in \Pi[a,b]$, then we say that f is of **bounded variation** *on $[a,b]$. The set of all functions of bounded variation on the interval $[a,b]$ is denoted by the symbol $BV[a,b]$.*

Comment 3.3.1.

1. *Note saying a function f is of bounded variation is equivalent to saying the set $\{\sum_{\pi} |\Delta f_j| \, : \, \pi \in \Pi[a,b]\}$ is bounded, and, therefore, has a supremum.*
2. *Also, if f is of bounded variation on $[a,b]$, then, for any $x \in (a,b)$, the set $\{a, x, b\}$ is a partition of $[a,b]$. Hence, there exists $M > 0$ such that $\mid f(x) - f(a) \mid + \mid f(b) - f(x) \mid \leq M$. But this implies*

$$\mid f(x) \mid - \mid f(a) \mid \quad \leq \quad \mid f(x) - f(a) \mid \, + \, \mid f(b) - f(x) \mid \leq \, M$$

This tells us that $| f(x) | \leq | f(a) | + M$. *Since our choice of* x *in* $[a, b]$ *was arbitrary, this shows that* f *is bounded, i.e.* $|| f ||_\infty < \infty$.

We can state the comments above formally as Theorem 3.3.1.

Theorem 3.3.1. Functions Of Bounded Variation Are Bounded

If f *is of bounded variation on* $[a, b]$, *then* f *is bounded on* $[a, b]$.

Theorem 3.3.2. Monotone Functions Are Of Bounded Variation

If f *is monotone on* $[a, b]$, *then* $f \in BV[a, b]$.

Proof. *As usual, we assume, for concreteness, that* f *is monotone increasing. Let* $\pi \in \Pi[a, b]$. *Hence, we can write* $\pi = \{x_0 = a, x_1, \ldots, x_{p-1}, x_p = b\}$. *Then*

$$\sum_{\pi} | \Delta f_j | = \sum_{\pi} | f(x_j) - f(x_{j-1}) | \,.$$

Since f *is monotone increasing, the absolute value signs are unnecessary, so that*

$$\sum_{\pi} | \Delta f_j | = \sum_{\pi} \Delta f_j = \sum_{\pi} \big(f(x_j) - f(x_{j-1}) \big).$$

But this is a telescoping sum, so

$$\sum_{\pi} \Delta f_j = f(x_p) - f(x_0) = f(b) - f(a).$$

Since the partition π *was arbitrary, it follows that* $\sum_\pi \Delta f_j \leq f(b) - f(a)$ *for all* $\pi \in \Pi[a, b]$. *This implies that* $f \in BV[a, b]$, *for if* $f(b) > f(a)$, *then we can simply let* $M = f(b) - f(a)$. *If* $f(b) = f(a)$, *then* f *must be constant, and we can let* $M = f(b) - f(a) + 1 = 1$. *In either case,* $f \in BV[a, b]$. ∎

Theorem 3.3.3. Bounded Differentiable Implies Bounded Variation

Suppose $f \in C[a, b]$, f *is differentiable on* (a, b), *and* $|| f' ||_\infty < \infty$. *Then* $f \in BV[a, b]$.

Proof. *Let* $\pi \in \Pi[a, b]$ *so that* $\pi = \{x_0 = a, x_1, \ldots, x_p = b\}$. *On each subinterval* $[x_{j-1}, x_j]$, *for* $1 \leq j \leq p$, *the hypotheses of the Mean Value Theorem are satisfied. Hence, there is a point,* $y_j \in (x_{j-1}, x_j)$, *with* $\Delta f_j = f(x_j) - f(x_{j-1}) = f'(y_j)\Delta x_j$. *So, we have*

$$| \Delta f_j | = | f'(y_j) | \Delta x_j \leq B \Delta x_j,$$

where B *is the bound on* f' *that we assume exists by hypothesis. Thus, for any* $\pi \in \Pi[a, b]$, *we have*

$$\sum_{\pi} | \Delta f_j | \leq B \sum_{\pi} \Delta x_j = B(b - a) < \infty.$$

Therefore, $f \in BV[a, b]$. ∎

Definition 3.3.2. The Total Variation Of A Function Of Bounded Variation
Let $f \in BV[a,b]$. The real number

$$V(f;a,b) = \sup\left\{\sum_{\pi} \mid \Delta f_j \mid : \ \pi \in \Pi[a,b]\right\}$$

is called the **Total Variation** *of f on $[a,b]$.*

Note that this number always exists if $f \in BV[a,b]$.

Comment 3.3.2. *For any $f \in BV[a,b]$, we clearly have $V(f;a,b) = V(-f;a,b)$ and $V(f;a,b) \geq 0$. Moreover, we also see that $V(f;a,b) = 0$ if and only if f is constant on $[a,b]$.*

Theorem 3.3.4. Functions Of Bounded Variation Are Closed Under Addition
If f and g are in $BV[a,b]$, then so are $f \pm g$, and $V(f \pm g;a,b) \leq V(f;a,b) + V(g;a,b)$.

Proof. *Let $\pi \in \Pi[a,b]$, so that $\pi = \{x_0 = a, x_1, \ldots, x_p = b\}$. Consider $f+g$ first. We have, for each $1 \leq j \leq p$,*

$$\begin{aligned} \mid \Delta(f+g)_j \mid &= \mid (f+g)(x_j) - (f+g)(x_{j-1}) \mid \\ &\leq \mid f(x_j) - f(x_{j-1}) \mid + \mid g(x_j) - g(x_{j-1}) \mid \\ &\leq \mid \Delta f_j \mid + \mid \Delta g_j \mid . \end{aligned}$$

This implies that, for any $\pi \in \Pi[a,b]$,

$$\sum_{\pi} \mid \Delta(f+g)_j \mid \leq \sum_{\pi} \mid \Delta f_j \mid + \sum_{\pi} \mid \Delta g_j \mid .$$

Both quantities on the right-hand side are bounded by $V(f;a,b)$ and $V(g;a,b)$, respectively. Since $\pi \in \Pi[a,b]$ was arbitrary, we have

$$V(f+g;a,b) \leq V(f;a,b) + V(g;a,b).$$

This shows that $f+g \in BV[a,b]$ and proves the desired inequality for that case. Since $V(-g;a,b) = V(g;a,b)$, we also have

$$V(f-g;a,b) \leq V(f;a,b) + V(-g;a,b) = V(f;a,b) + V(g;a,b),$$

which proves that $f-g \in BV[a,b]$. ■

Theorem 3.3.5. Products Of Functions Of Bounded Variation Are Of Bounded Variation
If $f,g \in BV[a,b]$, then $fg \in BV[a,b]$ and $V(fg;a,b) \leq \parallel g \parallel_\infty V(f;a,b) + \parallel f \parallel_\infty V(g;a,b)$.

Proof. *By Theorem 3.3.1, we know that f and g are bounded. Hence, the numbers $\| f \|_\infty$ and $\| g \|_\infty$ exist and are finite. Let $h = fg$, and let $\pi = \{x_0 = a, x_1, \ldots, x_p = b\}$ be any partition. Then*

$$\begin{aligned} |\, \Delta h_j \,| &= |\, f(x_j)g(x_j) - f(x_{j-1})g(x_{j-1}) \,| \\ &= |\, f(x_j)g(x_j) - g(x_j)f(x_{j-1}) + g(x_j)f(x_{j-1}) - f(x_{j-1})g(x_{j-1}) \,| \\ &\leq |\, g(x_j) \,||\, \Delta f_j \,| + |\, f(x_{j-1}) \,||\, \Delta g_j \,| \\ &\leq \| g \|_\infty |\, \Delta f_j \,| \; + \; \| f \|_\infty |\, \Delta g_j \,| \end{aligned}$$

Thus,

$$\begin{aligned} \sum_\pi |\, \Delta h_j \,| &\leq \| g \|_\infty \sum_\pi |\, \Delta f_j \,| + \| f \|_\infty \sum_\pi |\, \Delta g_j \,| \\ &\leq \| g \|_\infty \; V(f; a, b) + \| f \|_\infty \; V(g; a, b) \end{aligned}$$

Since π was arbitrary, we see the right hand side is an upper bound for all the partition sums and hence, the supremum of all these sums must also be less than or equal to the right hand side. Thus,

$$V(fg; a, b) \leq \| g \|_\infty \; V(f; a, b) + \| f \|_\infty \; V(g; a, b)$$

■

Comment 3.3.3. *Note that we have verified that $BV[a, b]$ is a* commutative algebra *(i.e. a* ring*) of functions with an identity, since the constant function $f = 1$ is of bounded variation.*

It is natural to ask, then, what the units are in this algebra. That is, what functions have multiplicative inverses?

Theorem 3.3.6. Inverses Of Functions Of Bounded Variation
Let f be in $BV[a, b]$, and assume that there is a positive m such that $|\, f(x) \,| \geq m > 0$ for all $x \in [a, b]$. Then $1/f \in BV[a, b]$ and $V(1/f; a, b) \leq (1/m^2)V(f; a, b)$.

Proof. *Let $\pi = \{x_0 = a, x_1, \ldots, x_p\}$ be any partition. Then*

$$\begin{aligned} \left| \Delta \Big(\frac{1}{f}\Big)_j \right| &= \left| \frac{1}{f(x_j)} - \frac{1}{f(x_{j-1})} \right| \\ &= \left| \frac{f(x_{j-1}) - f(x_j)}{f(x_j)f(x_{j-1})} \right| \\ &= \frac{|\, \Delta f_j \,|}{|\, f(x_j) \,||\, f(x_{j-1}) \,|} \\ &\leq \frac{\Delta f_j}{m^2}. \end{aligned}$$

Thus, we have

$$\sum_\pi \mid \Delta\left(\frac{1}{f}\right)_j \mid \quad \leq \quad \frac{1}{m^2}\sum_\pi \mid \Delta f_j \mid$$

implying that $V(1/f;a,b) \leq (1/m^2)V(f;a,b)$. ■

Comment 3.3.4.

1. *Any polynomial, p, is in* $BV[a,b]$*, and p is a unit if none of its zeros occur in the interval.*
2. *Any rational function* p/q *where p and q are of bounded variation on* $[a,b]$*, is in* $BV[a,b]$ *as long as none of the zeros of q occur in the interval.*
3. $e^x \in BV[a,b]$*. In fact,* $e^{u(x)} \in BV[a,b]$ *if* $u(x)$ *is monotone or has a bounded derivative.*
4. $\sin x$ *and* $\cos x$ *are in* $BV[a,b]$ *by Theorem 3.3.3.*
5. $\tan x \in BV[a,b]$ *if* $[a,b]$ *does not contain any point of the form* $(2k+1)\pi/2$ *for* $k \in \boldsymbol{Z}$*, by Theorem 3.3.6.*
6. *The function*

$$f(x) = \begin{cases} \sin\frac{1}{x}, & 0 < x \leq 1 \\ 0, & x = 0 \end{cases}$$

is not in $BV[0,1]$*. To see this, choose partition points* $\{x_0,\ldots,x_p\}$ *by* $x_0 = 0$*,* $x_p = 1$*, and*

$$x_j = \frac{2}{\pi(2p-2j+1)}, \qquad 1 \leq j \leq p-1.$$

Then

$$\Delta f_1 = \sin\left(\frac{\pi(2p-1)}{2}\right) = \pm 1,$$

$$\Delta f_2 = \sin\left(\frac{\pi(2p-3)}{2}\right) - \sin\left(\frac{\pi(2p-1)}{2}\right) = \pm 2,$$

and continuing, we find

$$\Delta f_{p-1} = \sin\left(\frac{\pi(2p-2(p-1)+1)}{2}\right) - \sin\left(\frac{\pi(2p-2(p-2)+1)}{2}\right) = \pm 2,$$

$$\Delta f_p = \sin 1 - \sin(3\pi/2) = \sin 1 + 1.$$

Thus,

$$\begin{aligned}\sum_\pi \quad &= \quad \mid \Delta f_1 \mid \, + \sum_{j=2}^{p-1} \mid \Delta f_j \mid \; + \sin 1 + 1 \\ &= \quad 2(p-1) \; + \; \sin 1.\end{aligned}$$

Hence, we can make the value of this sum as large as we desire and so this function is not of bounded variation.

3.3.1 Homework

Exercise 3.3.1. *Prove that if f is of bounded variation on the finite interval $[a, b]$, then αf is also of bounded variation for any scalar α. Do this proof using the partition approach.*

Exercise 3.3.2. *Prove that if f and g are of bounded variation on the finite interval $[a, b]$, then $\alpha f + \beta g$ is also of bounded variation for any scalars α and β. Do this proof using the partition approach. Note, these two exercises essentially show $BV[a, b]$ is a vector space.*

Exercise 3.3.3. *Prove $BV[a, b]$ is a complete normed linear space with norm $|| \cdot ||$ defined by*

$$|| f || \quad = \quad | f(a) | + V(f, a, b)$$

Exercise 3.3.4. *Define f on $[0, 1]$ by*

$$f(x) \quad = \quad \begin{cases} x^2 \cos(x^{-2}) & x \neq 0 \in [0, 1] \\ 0 & x = 0 \end{cases}$$

Prove that f is differentiable on $[0, 1]$ but is not of bounded variation. This is a nice example of something we will see later. This f is a function which is continuous but not absolutely continuous.

3.4 The Total Variation Function

Theorem 3.4.1. The Total Variation Is Additive On Intervals
If $f \in BV[a, b]$ and $c \in [a, b]$, then $f \in BV[a, c]$, $f \in BV[c, b]$, and $V(f; a, b) = V(f; a, c) + V(f; c, b)$. That is, the total variation, V, is additive on intervals.

Proof. *The case $c = a$ or $c = b$ is easy, so we assume $c \in (a, b)$. Let $\pi_1 \in \Pi[a, c]$ and $\pi_2 \in \Pi[c, b]$ with $\pi_1 = \{x_0 = a, x_1, \ldots, x_p = c\}$ and $\pi_2 = \{y_0 = c, y_1, \ldots, y_q = b\}$. Then $\pi_1 \vee \pi_2$ is a partition of $[a, b]$ and we know*

$$\sum_{\pi_1 \vee \pi_2} | \Delta f_j | = \sum_{\pi_1} | \Delta f_j | + \sum_{\pi_2} | \Delta f_j | \leq V(f; a, b).$$

Dropping the π_2 term, and noting that $\pi_1 \in \Pi[a, c]$ was arbitrary, we see that

$$\sup_{\pi_1 \in \Pi[a,c]} \sum_{\pi_1} | \Delta f_j | \leq V(f; a, b),$$

which implies that $V(f; a, c) \leq V(f; a, b) < \infty$. Thus, $f \in BV[a, c]$. A similar argument shows that $V(f; c, b) \leq V(f; a, b)$, so $f \in BV[c, b]$.

Finally, since both π_1 and π_2 were arbitrary and we know that

$$\sum_{\pi_1} | \Delta f_j | + \sum_{\pi_2} | \Delta f_j | \leq V(f; a, b),$$

we see that $V(f;a,c)+V(f;c,b)\le V(f;a,b)$.

Now we will establish the reverse inequality. Let $\pi\in\Pi[a,b]$, *so that* $\pi=\{x_0=a,x_1,\dots,x_p=b\}$. *First, assume that* c *is a partition point of* π, *so that* $c=x_{k_0}$ *for some* k_0. *Thus,* $\pi=\{x_0,\dots,x_{k_0}\}\cup\{x_{k_0},\dots,x_p\}$. *Let* $\pi_1=\{x_0,\dots,x_{k_0}\}\in\Pi[a,c]$ *and let* $\pi_2=\{x_{k_0},\dots,x_p\}\in\Pi[c,b]$. *From the first part of our proof, we know that* $f\in BV[a,c]$ *and* $f\in BV[c,b]$, *so*

$$\begin{aligned}\sum_{\pi}\mid\Delta f_j\mid &= \sum_{\pi_1}\mid\Delta f_j\mid+\sum_{\pi_2}\mid\Delta f_j\mid\\ &\le V(f;a,c)+V(f;c,b).\end{aligned}$$

Since $\pi\in\Pi[a,b]$ *was arbitrary, it follows that* $V(f;a,b)\le V(f;a,c)+V(f;c,b)$. *For the other case, suppose* c *is not a partition point of* π. *Then* c *must lie inside one of the subintervals. That is,* $c\in(x_{k_0-1},x_{k_0})$ *for some* k_0. *Let* $\pi'=\{x_0,\dots,x_{k_0-1},c,x_{k_0},\dots,x_p\}$ *be a new partition of* $[a,b]$. *Then* π' *refines* π. *Apply our previous argument to conclude that*

$$\sum_{\pi'}\mid\Delta f_j\mid\le V(f;a,c)+V(f;c,b).$$

Finally, we note that

$$\sum_{\pi}\mid\Delta f_j\mid\le\sum_{\pi'}\mid\Delta f_j\mid,$$

since

$$\mid f(x_{k_0})-f(x_{k_0-1})\mid\le\mid f(x_{k_0})-f(c)\mid+\mid f(c)-f(x_{k_0-1})\mid.$$

Thus, we have

$$\sum_{\pi}\mid\Delta f_j\mid\le V(f;a,c)+V(f;c,b).$$

Since π *was arbitrary, it follows that* $V(f;a,b)\le V(f;a,c)+V(f;c,b)$. *Combining these two inequalities, we see the result is established.* ■

Definition 3.4.1. The Variation Function Of a Function f Of Bounded Variation

Let $f\in BV[a,b]$. *The* **Variation Function** *of* f, *or simply the* **Variation** *of* f, *is the function* $V_f:[a,b]\to\Re$ *defined by*

$$V_f(x)=\begin{cases}0, & x=a\\ V(f;a,x), & a<x\le b\end{cases}$$

Theorem 3.4.2. V_f and V_f-f Are Monotone For a Function f of Bounded Variation

If $f\in BV[a,b]$, *then the functions* V_f *and* V_f-f *are monotone increasing on* $[a,b]$.

Proof. *Pick* $x_1,x_2\in[a,b]$ *with* $x_1<x_2$. *By Theorem 3.4.1,* $f\in BV[a,x_1]$ *and* $f\in BV[a,x_2]$. *Apply this same theorem to the interval* $[a,x_1]\cup[x_1,x_2]$ *to conclude that* $f\in BV[x_1,x_2]$. *Thus*

$$V_f(x_2) \;=\; V(f;a,x_2) \;=\; V(f;a,x_1)+V(f;x_1,x_2) \;=\; V_f(x_1)+V(f;x_1,x_2).$$

It follows that $V_f(x_2) - V_f(x_1) = V(f;x_1,x_2) \geq 0$, *so* V_f *is monotone increasing. Now, consider* $(V_f - f)(x_2) - (V_f - f)(x_1)$. *We have*

$$\begin{aligned}
(V_f - f)(x_2) - (V_f - f)(x_1) &= V_f(x_2) - V_f(x_1) - (f(x_2) - f(x_1)) \\
&= V(f;a,x_2) - V(f;a,x_1) - (f(x_2) - f(x_1)) \\
&= V(f;x_1,x_2) - (f(x_2) - f(x_1)).
\end{aligned}$$

But $\{x_1, x_2\}$ *is the trivial partition of* $[x_1, x_2]$, *so*

$$\sum_{\{x_1,x_2\}} \mid \Delta f_j \mid \;\leq\; \sup_{\pi \in \Pi[x_1,x_2]} \sum_{\pi} \mid \Delta f_j \mid = V(f;x_1,x_2).$$

Thus, $V(f;x_1,x_2) - (f(x_2) - f(x_1)) \geq 0$, *implying that* $V_f - f$ *is monotone increasing.* ∎

Theorem 3.4.3. A Function Of Bounded Variation Is The Difference of Two Increasing Functions
Every $f \in BV[a,b]$ *can be written as the difference of two monotone increasing functions on* $[a,b]$. *In other words,*

$$BV[a,b] \;=\; \{f : [a,b] \rightarrow \Re \mid \exists u, v : [a,b] \rightarrow \Re,\ u, v \text{ monotone increasing},\ f = u - v\}.$$

Proof. *If* $f = u - v$, *where* u *and* v *are monotone increasing, then* u *and* v *are of bounded variation. Since* $BV[a,b]$ *is an algebra, it follows that* $f \in BV[a,b]$.

Conversely, suppose $f \in BV[a,b]$, *and let* $u = V_f$ *and* $v = V_f - f$. *Then* u *and* v *are monotone increasing and* $u - v = f$. ∎

Comment 3.4.1. *Theorem 3.4.3 tells us if* g *is of bounded variation on* $[a,b]$, *then* $g = u - v$ *where* u *and* v *are monotone increasing. Thus, we can also use the Saltus decomposition of* u *and* v *to conclude*

$$\begin{aligned}
f &= (u_c + S_u) \;-\; (v_c + S_v) \\
&= (u_c - v_c) \;+\; (S_u - S_v)
\end{aligned}$$

The first term is the difference of two continuous functions of bounded variation and the second term is the difference of Saltus functions. This is essentially another form of decomposition theorem for a function of bounded variation.

3.5 Continuous Functions of Bounded Variation

Theorem 3.5.1. Functions Of Bounded Variation Always Possess Right and Left Hand Limits
Let $f \in BV[a,b]$. Then the limit $f(x^+)$ exists for all $x \in [a,b)$ and the limit $f(x^-)$ exists for all $x \in (a,b]$.

Proof. *By Theorem 3.4.2, V_f and $V_f - f$ are monotone increasing. So $V_f(x^+)$ and $(V_f - f)(x^+)$ both exist. Hence,*

$$\begin{aligned} f(x^+) &= \lim_{x \to x^+} f(x) \\ &= \lim_{x \to x^+} [V_f(x) - (V_f - f)(x)] \\ &= \lim_{x \to x^+} V_f(x) + \lim_{x \to x^+} (V_f - f)(x) \\ &= V_f(x^+) + (V_f - f)(x^+). \end{aligned}$$

So, $f(x^+)$ exists. A similar argument shows that $f(x^-)$ exists. ∎

Theorem 3.5.2. Functions Of Bounded Variation Have Countable Discontinuity Sets
If $f \in BV[a,b]$, then the set of discontinuities of f is countable.

Proof. *$f = u - v$ where u and v are monotone increasing. By Theorem 3.4.3, $S_1 = \{x \in [a,b] | u \text{is not continuous at} x\}$ and $S_2 = \{x \in [a,b] \mid v \text{is not continuous at} x\}$ are countable. The union of these sets is the set of all the points of possible discontinuity of f, so the set of discontinuities of f is countable.* ∎

Theorem 3.5.3. $f \in BV[a,b]$ Is Continuous If and Only If V_f Is Continuous
Let $f \in BV[a,b]$. Then f is continuous at $c \in [a,b]$ if and only if V_f is continuous at c.

Proof. *The case where $c = a$ and $c = b$ are easier, so we will only prove the case where $c \in (a,b)$. First, suppose f is continuous at c. We will prove separately that V_f is continuous from the right at c and from the left at c.*

Let $\epsilon > 0$ be given. Since f is continuous at c, there is a positive δ such that if x is in $(c-\delta, c+\delta) \subset [a,b]$, then $| f(x) - f(c) | < \epsilon/2$. Now,

$$V(f;c,b) = \sup_{\pi \in \Pi[c,b]} \sum_{\pi} | \Delta f_j | .$$

So, there is a partition π_0 such that

$$V(f;c,b) - \frac{\epsilon}{2} \quad < \quad \sum_{\pi_0} | \Delta f_j | \leq V(f;c,b). \tag{$*$}$$

If $\pi_0{}'$ is any refinement of π_0, we see that

$$\sum_{\pi_0} \mid \Delta f_j \mid \leq \sum_{\pi_0{}'} \mid \Delta f_j \mid,$$

since adding points to π_0 simply increases the sum. Thus,

$$V(f;c,b) - \frac{\epsilon}{2} \quad < \quad \sum_{\pi_0} \mid \Delta f_j \mid \leq \sum_{\pi_0{}'} \mid \Delta f_j \mid \leq V(f;c,b)$$

for any refinement $\pi_0{}'$ of π_0. Now, choose a partition, π_1 which refines π_0 and satisfies $\| \pi_1 \| < \delta$. Then

$$V(f;c,b) - \frac{\epsilon}{2} \quad < \quad \sum_{\pi_1} \mid \Delta f_j \mid \leq V(f;c,b). \tag{$**$}$$

So, if $\pi_1 = \{x_0 = c, x_1, \ldots, x_p\}$, then $\mid x_1 - x_0 \mid < \delta$. Thus, we have $\mid x_1 - c \mid < \delta$. It follows that $\mid f(x_1) - f(c) \mid < \epsilon/2$. From Equation $**$, we then have

$$\begin{aligned}
V(f;c,b) - \frac{\epsilon}{2} \quad &< \quad \sum_{\pi_1} \mid \Delta f_j \mid \\
&= \quad \mid f(x_1) - f(c) \mid + \sum_{rest\ of\ \pi_1} \mid \Delta f_j \mid \\
&< \quad \frac{\epsilon}{2} + \sum_{rest\ of\ \pi_1} \mid \Delta f_j \mid \\
&< \quad \frac{\epsilon}{2} + V(f;x_1,b).
\end{aligned}$$

So, we see that

$$V(f;c,b) - \frac{\epsilon}{2} < \frac{\epsilon}{2} + V(f;x_1,b),$$

which implies that

$$V(f;c,b) - V(f;x_1,b) \quad < \quad \frac{\epsilon}{2} + \frac{\epsilon}{2} \; = \; \epsilon.$$

But $V(f;c,b) - V(f;x_1,b) = V(f;c,x_1)$ which is the same as $V_f(x_1) - V_f(c)$. Thus, we have

$$V_f(x_1) - V_f(c) \quad < \quad \epsilon.$$

Now V_f is monotone and hence we have shown that if $x \in (c, x_1)$,

$$V_f(x) - V_f(c) \quad \leq \quad V_f(x_1) - V_f(c) \; < \; \epsilon.$$

Since $\epsilon > 0$ was arbitrary, this verifies the right continuity of V_f at c.

The argument for left continuity is similar. We can find a partition π_1 *of* $[a, c]$ *with partition points* $\{x_0 = a, x_1, \ldots, x_{p-1}, x_p = c\}$ *such that* $|| \pi_1 || < \delta$ *and*

$$\begin{aligned} V(f; a, c) - \frac{\epsilon}{2} &< \mid f(c) - f(x_{p-1}) \mid + \sum_{rest\ of\ \pi_1} \mid \Delta f_j \mid \\ &\leq \mid f(c) - f(x_{p-1}) \mid + V(f; a, x_{p-1}). \end{aligned}$$

Since $|| \pi_1 || < \delta$*, we see as before that* $\mid f(c) - f(x_{p-1}) \mid < \epsilon/2$*. Thus,*

$$V(f; a, c) - \frac{\epsilon}{2} < \frac{\epsilon}{2} + V(f; a, x_{p-1}),$$

and it follows that

$$V(f; a, c) - V(f; a, x_{p-1}) < \epsilon,$$

or

$$V_f(c) - V_f(x_{p-1}) < \epsilon.$$

Since V_f *is monotone, we then have for any* x *in* (x_{p-1}, c) *that*

$$V_f(c) - V_f(x) < V_f(c) - V_f(x_{p-1}) < \epsilon$$

which shows the left continuity of V_f *at* c*. Hence,* V_f *is continuous at* c*.*

Conversely, suppose V_f *is continuous at* $c \in (a, b)$*. Given* $\epsilon > 0$*, there is a positive* δ *such that* $(c - \delta, c + \delta) \subset [a, b]$ *and* $\mid V_f(x) - V_f(c) \mid < \epsilon$ *for all* $x \in (c - \delta, c + \delta)$*. Pick any* $x \in (c, c + \delta)$*. Then* $\{c, x\}$ *is a trivial partition of* $[c, x]$*. Hence*

$$0 \leq \mid f(x) - f(c) \mid \leq V(f; c, x) = V(f; a, x) - V(f; a, c)$$

or

$$0 \leq \mid f(x) - f(c) \mid \leq V_f(x) - V_f(c) < \epsilon.$$

Hence, it follows that f *is continuous from the right. A similar argument shows that* f *is continuous from the left.* ■

We immediately have this corollary.

Theorem 3.5.4. $f \in BV[a, b] \cap C[a, b]$ **If and Only If** V_f **and** $V_f - f$ **Are Continuous and Increasing**

$f \in C[a, b] \cap BV[a, b]$ *if and only if* V_f *and* $V_f - f$ *are monotone increasing and continuous.*

Chapter 4

The Theory Of Riemann Integration

We will now develop the theory of the Riemann Integral for a bounded function f on the interval $[a, b]$. We followed the development of this material in (Fulks (3) 1978) closely at times, although Fulks does not cover some of the sections very well.

4.1 Defining The Riemann Integral

Definition 4.1.1. The Riemann Sum

Let $f \in B[a, b]$, and let $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]$ be given by $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$. Define $\boldsymbol{\sigma} = \{s_1, \ldots, s_p\}$, where $s_j \in [x_{j-1}, x_j]$ for $1 \leq j \leq p$. We call $\boldsymbol{\sigma}$ an evaluation set, *and we denote this by $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$. The* **Riemann Sum** *determined by the partition $\boldsymbol{\pi}$ and the evaluation set $\boldsymbol{\sigma}$ is defined by*

$$S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) = \sum_{\boldsymbol{\pi}} f(s_j) \Delta x_j$$

Definition 4.1.2. Riemann Integrability Of a Bounded f

We say $f \in B[a, b]$ is **Riemann Integrable** *on $[a, b]$ if there exists a real number, I, such that for every $\epsilon > 0$ there is a partition, $\boldsymbol{\pi_0} \in \boldsymbol{\Pi}[a, b]$ such that*

$$\mid S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I \mid < \epsilon$$

for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ and any evaluation set, $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$. We denote this value, I, by

$$I \equiv RI(f; a, b)$$

We denote the set of Riemann integrable functions on $[a, b]$ by $RI[a, b]$. Also, it is readily seen that the number $RI(f; a, b)$ in the definition above, when it exists, is unique. So we can speak of *Riemann Integral* of a function, f. We also have the following conventions.

1. $RI(f; a, b) \;=\; -RI(f; b, a)$
2. $RI(f; a, a) \;=\; 0$
3. f is called the **integrand**.

Theorem 4.1.1. $RI[a, b]$ **Is A Vector Space and** $RI(f; a, b)$ **Is A Linear Mapping**
$RI[a, b]$ is a vector space over $\Re$ and the mapping $I_R : RI[a, b] \rightarrow \Re$ defined by

$$I_R(f) = RI(f; a, b)$$

is a linear mapping.

Proof. *Let $f_1, f_2 \in RI[a, b]$, and let $\alpha, \beta \in \Re$. For any $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]$ and $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, we have*

$$\begin{aligned} S(\alpha f_1 + \beta f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) &= \sum_{\boldsymbol{\pi}} (\alpha f_1 + \beta f_2)(s_j) \Delta x_j \\ &= \alpha \sum_{\boldsymbol{\pi}} f_1(s_j) \Delta x_j + \beta \sum_{\boldsymbol{\pi}} f_2(s_j) \Delta x_j \\ &= \alpha S(f_1, \boldsymbol{\pi}, \boldsymbol{\sigma}) + \beta S(f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}). \end{aligned}$$

Since f_1 is Riemann integrable, given $\epsilon > 0$, there is a real number $I_1 = RI(f_1, a, b)$ and a partition $\boldsymbol{\pi}_1 \in \boldsymbol{\Pi}[a, b]$ such that

$$\mid S(f_1, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_1 \mid \;<\; \frac{\epsilon}{2(\mid \alpha \mid + 1)} \qquad (*)$$

for all refinements, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_1$, and all $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$.
Likewise, since f_2 is Riemann integrable, there is a real number $I_2 = RE(f_2; a, b)$ and a partition $\boldsymbol{\pi}_2 \in \boldsymbol{\Pi}[a, b]$ such that

$$\mid S(f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_2 \mid \;<\; \frac{\epsilon}{2(\mid \beta \mid + 1)} \qquad (**)$$

for all refinements, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_2$, and all $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$.
Let $\boldsymbol{\pi}_0 = \boldsymbol{\pi}_1 \vee \boldsymbol{\pi}_2$. Then $\boldsymbol{\pi}_0$ is a refinement of both $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$. So, for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$, and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, we have Equation $*$ and Equation $**$ are valid. Hence,

$$\begin{aligned} \mid S(f_1, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_1 \mid &< \frac{\epsilon}{2(\mid \alpha \mid + 1)} \\ \mid S(f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_2 \mid &< \frac{\epsilon}{2(\mid \beta \mid + 1)}. \end{aligned}$$

Thus, for any refinement $\boldsymbol{\pi}$ of $\boldsymbol{\pi}_0$ and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, it follows that

$$\begin{aligned}
\mid S(\alpha f_1 + \beta f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) - (\alpha I_1 + \beta I_2) \mid &= \mid \alpha S(f_1, \boldsymbol{\pi}, \boldsymbol{\sigma}) + \beta S(f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) - \alpha I_1 - \beta I_2 \mid \\
&\leq \mid \alpha \mid\mid S(f_1, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_1 \mid + \mid \beta \mid\mid S(f_2, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I_2 \mid \\
&< \mid \alpha \mid \frac{\epsilon}{2(\mid \alpha \mid +1)} + \mid \beta \mid \frac{\epsilon}{2(\mid \beta \mid +1)} \\
&< \epsilon.
\end{aligned}$$

This shows that $\alpha f_1 + \beta f_2$ is Riemann integrable and that the value of the integral $RI(\alpha f_1 + \beta f_2; a, b)$ is given by $\alpha RI(f_1; a, b) + \beta RI(f_2; a, b)$. It then follows immediately that I_R is a linear mapping.
■

Theorem 4.1.2. Fundamental Riemann Integral Estimates
Let $f \in RI[a, b]$. Let $m = \inf_x f(x)$ and let $M = \sup_x f(x)$. Then

$$m(b-a) \leq RI(f; a, b) \leq M(b-a).$$

Proof. *If $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]$, then for all $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, we see that*

$$\sum_{\boldsymbol{\pi}} m\Delta x_j \leq \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j \leq \sum_{\boldsymbol{\pi}} M\Delta x_j.$$

But $\sum_{\boldsymbol{\pi}} \Delta x_j = b - a$, so

$$m(b-a) \leq \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j \leq M(b-a),$$

or

$$m(b-a) \leq S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \leq M(b-a),$$

for any partition $\boldsymbol{\pi}$ and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$.

Now, let $\epsilon > 0$ be given. Then there exist $\boldsymbol{\pi}_0 \in \boldsymbol{\Pi}[a, b]$ such that for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$,

$$RI(f; a, b) - \epsilon < S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) < RI(f; a, b) + \epsilon.$$

Hence, for any such refinement, $\boldsymbol{\pi}$, and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, we have

$$m(b-a) \leq S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) < RI(f; a, b) + \epsilon$$

and

$$M(b-a) \geq S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) > RI(f; a, b) - \epsilon.$$

Since $\epsilon > 0$ is arbitrary, it follows that

$$m(b-a) \leq RI(f; a, b) \leq M(b-a).$$

■

Theorem 4.1.3. The Riemann Integral Is Order Preserving
The Riemann integral is order preserving. That is, if $f, f_1, f_2 \in RI[a,b]$, then

(i)
$$f \geq 0 \Rightarrow RI(f; a, b) \geq 0;$$

(ii)
$$f_1 \leq f_2 \Rightarrow RI(f_1; a, b) \leq RI(f_2; a, b).$$

Proof. *If $f \geq 0$ on $[a,b]$, then $\inf_x f(x) = m \geq 0$. Hence, by Theorem 4.1.2*

$$\int_a^b f(x)dx \geq m(b-a) \geq 0.$$

This proves the first assertion. To prove (ii), let $f = f_2 - f_1$. Then $f \geq 0$, and the second result follows from the first. ■

4.2 The Existence of the Riemann Integral: Darboux Integration

Although we have a definition for what it means for a bounded function to be Riemann integrable, we still do not actually know that $RI[a,b]$ is nonempty! In this section, we will show how we prove that the set of Riemann integrable functions is quite rich and varied.

Definition 4.2.1. Darboux Upper and Lower Sums
Let $f \in B[a,b]$. Let $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]$ be given by $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$. Define

$$m_j = \inf_{x_{j-1} \leq x \leq x_j} f(x) \qquad 1 \leq j \leq p,$$

and

$$M_j = \sup_{x_{j-1} \leq x \leq x_j} f(x) \qquad 1 \leq j \leq p.$$

We define the **Lower Darboux Sum** *by*

$$L(f, \boldsymbol{\pi}) = \sum_{\boldsymbol{\pi}} m_j \Delta x_j$$

and the **Upper Darboux Sum** *by*

$$U(f, \boldsymbol{\pi}) = \sum_{\boldsymbol{\pi}} M_j \Delta x_j.$$

Comment 4.2.1.

1. *It is straightforward to see that*

$$L(f,\boldsymbol{\pi}) \;\leq\; S(f,\boldsymbol{\pi},\boldsymbol{\sigma}) \leq U(f,\boldsymbol{\pi})$$

for all $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]$.

2. *We also have*

$$U(f,\boldsymbol{\pi}) - L(f,\boldsymbol{\pi}) \;=\; \sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j.$$

Theorem 4.2.1. $\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$ **Implies** $L(f,\boldsymbol{\pi}) \leq L(f,\boldsymbol{\pi}')$ **and** $U(f,\boldsymbol{\pi}) \geq U(f,\boldsymbol{\pi}')$

If $\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$*, that is, if* $\boldsymbol{\pi}'$ *refines* $\boldsymbol{\pi}$*, then* $L(f,\boldsymbol{\pi}) \leq L(f,\boldsymbol{\pi}')$ *and* $U(f,\boldsymbol{\pi}) \geq U(f,\boldsymbol{\pi}')$*.*

Proof. *The general result is established by induction on the number of points added. It is actually quite an involved induction. Here are some of the details:*

Step 1 *We prove the proposition for inserting points* $\{z_1,\ldots,z_q\}$ *into one subinterval of* $\boldsymbol{\pi}$*. The argument consists of*

1. *The* Basis *Step where we prove the proposition for the insertion of a single point into one subinterval.*
2. *The* Induction *Step where we assume the proposition holds for the insertion of* q *points into one subinterval and then we show the proposition still holds if an additional point is inserted.*
3. *With the* Induction *Step verified, the Principle of Mathematical Induction then tells us that the proposition is true for any refinement of* $\boldsymbol{\pi}$ *which places points into one subinterval of* $\boldsymbol{\pi}$*.*

Basis:

Subproof. *Let* $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]$ *be given by* $\{x_0 = a, x_1, \ldots, x_p = b\}$*. Suppose we form the refinement,* $\boldsymbol{\pi}'$*, by adding a single point* x' *to* $\boldsymbol{\pi}$*. into the interior of the subinterval* $[x_{k_0-1}, x_{k_0}]$*. Let*

$$\begin{aligned} m' &= \inf_{[x_{k_0-1},x']} f(x) \\ m'' &= \inf_{[x',x_{k_0}]} f(x). \end{aligned}$$

Note that $m_{k_0} = \min\{m', m''\}$ *and*

$$\begin{aligned} m_{k_0}\Delta x_{k_0} &= m_{k_0}(x_{k_0} - x_{k_0-1}) \\ &= m_{k_0}(x_{k_0} - x') + m_{k_0}(x' - x_{k_0-1}) \\ &\leq m''(x_{k_0} - x') + m'(x' - x_{k_0-1}) \\ &\leq m''\Delta x'' + m'\Delta x', \end{aligned}$$

where $\Delta x'' = x_{k_0} - x'$ *and* $\Delta x' = x' - x_{k_0-1}$*. It follows that*

$$\begin{aligned} L(f, \boldsymbol{\pi}') &= \sum_{j \neq k_0} m_j \Delta x_j + m' \Delta x' + m'' \Delta x'' \\ &\geq \sum_{j \neq k_0} m_j \Delta x_j + m_{k_0} \Delta x_{k_0} \\ &\geq L(f, \boldsymbol{\pi}). \end{aligned}$$

□

Induction:

Subproof. *We assume that q points $\{z_1, \ldots, z_q\}$ have been inserted into the subinterval $[x_{k_0-1}, x_{k_0}]$. Let $\boldsymbol{\pi'}$ denote the resulting refinement of $\boldsymbol{\pi}$. We assume that*

$$L(f, \boldsymbol{\pi}) \quad \leq \quad L(f, \boldsymbol{\pi'})$$

let the additional point added to this subinterval be called x' and call $\boldsymbol{\pi''}$ the resulting refinement of $\boldsymbol{\pi'}$. We know that $\boldsymbol{\pi'}$ has broken $[x_{k_0-1}, x_{k_0}]$ into $q+1$ pieces. For convenience of notation, let's label these $q+1$ subintervals as $[y_{j-1}, y_j]$ where y_0 is x_{k_0-1} and y_{q+1} is x_{k_0} and the y_j values in between are the original z_i points for appropriate indices. The new point x' is thus added to one of these $q+1$ pieces, call it $[y_{j_0-1}, y_{j_0}]$ for some index j_0. This interval plays the role of the original subinterval in the proof of the em Basis Step. An argument similar to that in the proof of the Basis *Step then shows us that*

$$L(f, \boldsymbol{\pi'}) \quad \leq \quad L(f, \boldsymbol{\pi''})$$

Combining with the first inequality from the Induction *hypothesis, we establish the result. Thus, the* Induction *Step is proved.* □

Step 2 *Next, we allow the insertion of a finite number of points into a finite number of subintervals of $\boldsymbol{\pi}$. The induction is now on the number of subintervals.*

1. *The* Basis *Step where we prove the proposition for the insertion of points into one subinterval.*
2. *The* Induction *Step where we assume the proposition holds for the insertion of points into q subintervals and then we show the proposition still holds if an additional subinterval has points inserted.*
3. *With the* Induction *Step verified, the Principle of Mathematical Induction then tells us that the proposition is true for any refinement of $\boldsymbol{\pi}$ which places points into any number of subintervals of $\boldsymbol{\pi}$.*

Basis

Subproof. *Step* 1 *above gives us the* Basis *Step for this proposition.* □

Induction

Subproof. *We assume the results holds for p subintervals and show it also holds when one more subinterval is added. Specifically, let $\boldsymbol{\pi}'$ be the refinement that results from adding points to p subintervals of $\boldsymbol{\pi}$. Then the* Induction *hypothesis tells us that*

$$L(f, \boldsymbol{\pi}) \quad \leq \quad L(f, \boldsymbol{\pi}')$$

Let $\boldsymbol{\pi}''$ denote the new refinement of $\boldsymbol{\pi}$ which results from adding more points into one more subinterval of $\boldsymbol{\pi}$. Then $\boldsymbol{\pi}''$ is also a refinement of $\boldsymbol{\pi}'$ where all the new points are added to one subinterval of $\boldsymbol{\pi}'$. Thus, Step 1 holds for the pair ($\boldsymbol{\pi}'$,$\boldsymbol{\pi}''$). We see

$$L(f, \boldsymbol{\pi}') \quad \leq \quad L(f, \boldsymbol{\pi}'')$$

and the desired result follows immediately. □

A similar argument establishes the result for upper sums. ■

Theorem 4.2.2. $L(f, \boldsymbol{\pi}_1) \leq U(f, \boldsymbol{\pi}_2)$

Let $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ be any two partitions in $\boldsymbol{\Pi}[a, b]$. Then $L(f, \boldsymbol{\pi}_1) \leq U(f, \boldsymbol{\pi}_2)$.

Proof. *Let $\boldsymbol{\pi} = \boldsymbol{\pi}_1 \vee \boldsymbol{\pi}_2$ be the common refinement of $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$. Then, by the previous result, we have*

$$L(f, \boldsymbol{\pi}_1) \leq L(f, \boldsymbol{\pi}) \leq U(f, \boldsymbol{\pi}) \leq U(f, \boldsymbol{\pi}_2).$$

■

Theorem 4.2.2 then allows us to define a new type of integrability for the bounded function f. We begin by looking at the infimum of the upper sums and the supremum of the lower sums for a given bounded function f.

Theorem 4.2.3. The Upper And Lower Darboux Integral Are Finite

Let $f \in B[a, b]$. Let $\mathscr{L} = \{L(f, \boldsymbol{\pi}) \mid \boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]\}$ and $\mathscr{U} = \{U(f, \boldsymbol{\pi}) \mid \boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]\}$. Define $L(f) = \sup \mathscr{L}$, and $U(f) = \inf \mathscr{U}$. Then $L(f)$ and $U(f)$ are both finite. Moreover, $L(f) \leq U(f)$.

Proof. *By Theorem 4.2.2, the set $\mathscr{L}$ is bounded above by any upper sum for f. Hence, it has a finite supremum and so $\sup \mathscr{L}$ is finite. Also, again by Theorem 4.2.2, the set $\mathscr{U}$ is bounded below by any lower sum for f. Hence, $\inf \mathscr{U}$ is finite. Finally, since $L(f) \leq U(f, \boldsymbol{\pi})$ and $U(f) \geq L(f, \boldsymbol{\pi})$ for all $\boldsymbol{\pi}$, by definition of the infimum and supremum of a set of numbers, we must have $L(f) \leq U(f)$.* ■

Definition 4.2.2. Darboux Lower And Upper Integrals

Let f be in $B[a, b]$. The **Lower Darboux Integral** *of f is defined to be the finite number $L(f) = \sup \mathscr{L}$, and the* **Upper Darboux Integral** *of f is the finite number $U(f) = \inf \mathscr{U}$.*

We can then define what is meant by a bounded function being *Darboux Integrable* on $[a, b]$.

Definition 4.2.3. Darboux Integrability

Let f be in $B[a, b]$. We say f is Darboux Integrable on $[a, b]$ if $L(f) = U(f)$. The common value is then called the Darboux Integral of f on $[a, b]$ and is denoted by the symbol $DI(f; a, b)$.

Comment 4.2.2. *Not all bounded functions are Darboux Integrable. Consider the function $f : [0, 1] \to \Re$ defined by*

$$f(t) = \begin{cases} 1 & t \in [0,1] \text{ and is rational} \\ -1 & t \in [0,1] \text{ and is irrational} \end{cases}$$

You should be able to see that for any partition of $[0, 1]$, the infimum of f on any subinterval is always -1 as any subinterval contains irrational numbers. Similarly, any subinterval contains rational numbers and so the supremum of f on a subinterval is 1. Thus $U(f, \boldsymbol{\pi}) = 1$ and $L(f, \boldsymbol{\pi}) = -1$ for any partition $\boldsymbol{\pi}$ of $[0, 1]$. It follows that $L(f) = -1$ and $U(f) = 1$. Thus, f is bounded but not Darboux Integrable.

Definition 4.2.4. Riemann's Criterion for Integrability

Let $f \in B[a, b]$. We say that **Riemann's Criteria** *holds for f if for every positive ϵ there exists a $\boldsymbol{\pi}_0 \in \boldsymbol{\Pi}[a, b]$ such that $U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) < \epsilon$ for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$.*

Theorem 4.2.4. The Riemann Integral Equivalence Theorem

Let $f \in B[a, b]$. Then the following are equivalent.

(i) $f \in RI[a, b]$.

(ii) *f satisfies Riemann's Criteria.*

(iii) *f is Darboux Integrable, i.e, $L(f) = U(f)$, and $RI(f; a, b) = DI(f; a, b)$.*

Proof.
$(i) \Rightarrow (ii)$

Subproof. *Assume $f \in RI[a, b]$, and let $\epsilon > 0$ be given. Let IR be the Riemann integral of f over $[a, b]$. Choose $\boldsymbol{\pi}_0 \in \boldsymbol{\Pi}[a, b]$ such that $\mid S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) - IR \mid < \epsilon/3$ for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$. Let $\boldsymbol{\pi}$ be any such refinement, denoted by $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$, and let m_j, M_j be defined as usual. Using the Infimum and Supremum Tolerance Lemmas, we can conclude that, for each $j = 1, \ldots, p$, there exist $s_j, t_j \in [x_{j-1}, x_j]$ such that*

$$M_j - \frac{\epsilon}{6(b-a)} < f(s_j) \leq M_j$$

$$m_j \leq f(t_j) < m_j + \frac{\epsilon}{6(b-a)}.$$

It follows that

$$f(s_j) - f(t_j) > M_j - \frac{\epsilon}{6(b-a)} - m_j - \frac{\epsilon}{6(b-a)}.$$

Thus, we have

$$M_j - m_j - \frac{\epsilon}{3(b-a)} < f(s_j) - f(t_j).$$

Multiply this inequality by Δx_j to obtain

$$(M_j - m_j)\Delta x_j - \frac{\epsilon}{3(b-a)}\Delta x_j < \big(f(s_j) - f(t_j)\big)\Delta x_j.$$

Now, sum over $\boldsymbol{\pi}$ to obtain

$$\begin{aligned} U(f,\boldsymbol{\pi}) - L(f,\boldsymbol{\pi}) &= \sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j \\ &< \frac{\epsilon}{3(b-a)} \sum_{\boldsymbol{\pi}} \Delta x_j \; + \; \sum_{\boldsymbol{\pi}} \big(f(s_j) - f(t_j)\big)\Delta x_j. \end{aligned}$$

This simplifies to

$$\sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j - \frac{\epsilon}{3} \; < \; \sum_{\boldsymbol{\pi}} \big(f(s_j) - f(t_j)\big)\Delta x_j. \qquad (*)$$

Now, we have

$$\begin{aligned} \Big| \sum_{\boldsymbol{\pi}} \big(f(s_j) - f(t_j)\big)\Delta x_j \Big| &= \Big| \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j - \sum_{\boldsymbol{\pi}} f(t_j)\Delta x_j \Big| \\ &= \Big| \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j - IR + IR - \sum_{\boldsymbol{\pi}} f(t_j)\Delta x_j \Big| \\ &\leq \Big| \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j - IR \Big| + \Big| \sum_{\boldsymbol{\pi}} f(t_j)\Delta x_j - IR \Big| \\ &= | \, S(f,\boldsymbol{\pi},\boldsymbol{\sigma}_s) - IR \, | + | \, S(f,\boldsymbol{\pi},\boldsymbol{\sigma}_t) - IR \, |, \end{aligned}$$

where $\boldsymbol{\sigma}_s = \{s_1, \ldots, s_p\}$ and $\boldsymbol{\sigma}_t = \{t_1, \ldots, t_p\}$ are evaluation sets of $\boldsymbol{\pi}$. Now, by our choice of partition $\boldsymbol{\pi}$, we know

$$\begin{aligned} | \, S(f,\boldsymbol{\pi},\boldsymbol{\sigma}_s) - IR \, | &< \frac{\epsilon}{3} \\ | \, S(f,\boldsymbol{\pi},\boldsymbol{\sigma}_t) - IR \, | &< \frac{\epsilon}{3}. \end{aligned}$$

Thus, we can conclude that

$$\Big| \sum_{\boldsymbol{\pi}} \big(f(s_j) - f(t_j)\big)\Delta x_j \Big| < \frac{2\epsilon}{3}.$$

Applying this to the inequality in Equation $*$, we obtain

$$\sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j < \epsilon.$$

Now, $\boldsymbol{\pi}$ was an arbitrary refinement of $\boldsymbol{\pi}_0$, and $\epsilon > 0$ was also arbitrary. So this shows that f satisfies Riemann's condition. □

(ii) $\Rightarrow$ *(iii)*

Subproof. *Now, assume that f satisfies Riemann's criteria, and let $\epsilon > 0$ be given. Then there is a partition, $\boldsymbol{\pi}_0 \in \mathbf{\Pi}[a,b]$ such that $U(f,\boldsymbol{\pi}) - L(f,\boldsymbol{\pi}) < \epsilon$ for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$. Thus, by the definition of the upper and lower Darboux integrals, we have*

$$U(f) \leq U(f,\boldsymbol{\pi}) < L(f,\boldsymbol{\pi}) + \epsilon \leq L(f) + \epsilon.$$

Since ϵ is arbitrary, this shows that $U(f) \leq L(f)$. The reverse inequality has already been established. Thus, we see that $U(f) = L(f)$. □

(iii) $\Rightarrow$ *(i)*

Subproof. *Finally, assume f is Darboux integral which means $L(f) = U(f)$. Let ID denote the value of the Darboux integral. We will show that f is also Riemann integrable according to the definition and that the value of the integral is ID.*

Let $\epsilon > 0$ be given. Now, recall that

$$\begin{aligned} ID &= L(f) = \sup_{\boldsymbol{\pi}} L(f,\boldsymbol{\pi}) \\ &= U(f) = \inf_{\boldsymbol{\pi}} U(f,\boldsymbol{\pi}) \end{aligned}$$

Hence, by the Supremum Tolerance Lemma, there exists $\boldsymbol{\pi}_1 \in \mathbf{\Pi}[a,b]$ such that

$$ID - \epsilon = L(f) - \epsilon < L(f,\boldsymbol{\pi}_1) \leq L(f) = ID$$

and by the Infimum Tolerance Lemma, there exists $\boldsymbol{\pi}_2 \in \mathbf{\Pi}[a,b]$ such that

$$ID = U(f) \leq U(f,\boldsymbol{\pi}_2) < U(f) + \epsilon = ID + \epsilon.$$

Let $\boldsymbol{\pi}_0 = \boldsymbol{\pi}_1 \vee \boldsymbol{\pi}_2$ be the common refinement of $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$. Now, let $\boldsymbol{\pi}$ be any refinement of $\boldsymbol{\pi}_0$, and let $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$ be any evaluation set. Then we have

$$ID - \epsilon < L(f,\boldsymbol{\pi}_1) \leq L(f,\boldsymbol{\pi}_0) \leq L(f,\boldsymbol{\pi}) \leq S(f,\boldsymbol{\pi},\boldsymbol{\sigma}) \leq U(f,\boldsymbol{\pi}) \leq U(f,\boldsymbol{\pi}_0) \leq U(f,\boldsymbol{\pi}_2) < ID + \epsilon.$$

Thus, it follows that

$$ID - \epsilon < S(f,\boldsymbol{\pi},\boldsymbol{\sigma}) < ID + \epsilon.$$

Since the refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ was arbitrary, as were the evaluation set, $\boldsymbol{\sigma}$, and the tolerance ϵ, it follows that for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ and any $\epsilon > 0$, we have

$$| S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) - ID | < \epsilon.$$

This shows that f is Riemann Integrable and the value of the integral is ID. □

■

Comment 4.2.3. *By Theorem 4.2.4, we now know that the Darboux and Riemann integral are equivalent. Hence, it is now longer necessary to use a different notation for these two different approaches to what we call integration. From now on, we will use this notation*

$$RI(f; a, b) \equiv DI(f; a, b) \equiv \int f(t)\, dt$$

where the (t) in the new integration symbol refers to the name we wish to use for the independent variable and dt is a mnemonic to remind us that the $|| \boldsymbol{\pi} ||$ is approaching zero as we choose progressively finer partitions of $[a, b]$. This is, of course, not very rigorous notation. A better notation would be

$$RI(f; a, b) \equiv DI(f; a, b) \equiv I(f; a, b)$$

where the symbol I denotes that we are interested in computing the integral of f using the equivalent approach of Riemann or Darboux. Indeed, the notation $I(f; a, b)$ does not require the uncomfortable lack of rigor that the symbol dt implies. However, for historical reasons, the symbol $\int f(t)\, dt$ will be used.

Also, the use of the $\int f(t)\, dt$ allows us to very efficiently apply the integration techniques of substitution and so forth as we have shown in Chapter 2.

4.3 Properties Of The Riemann Integral

We can now prove a series of properties of the Riemann integral.

Theorem 4.3.1. Properties Of The Riemann Integral

Let $f, g \in RI[a,b]$. Then

(i) $\mid f \mid \in RI[a,b]$;

(ii)

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b \mid f \mid dx;$$

(iii) $f^+ = \max\{f, 0\} \in RI[a,b]$;

(iv) $f^- = \max\{-f, 0\} \in RI[a,b]$;

(v)

$$\begin{aligned} \int_a^b f(x)dx &= \int_a^b [f^+(x) - f^-(x)]dx = \int_a^b f^+(x)dx - \int_a^b f^-(x)dx \\ \int_a^b \mid f(x) \mid dx &= \int_a^b [f^+(x) + f^-(x)]dx = \int_a^b f^+(x)dx + \int_a^b f^-(x)dx; \end{aligned}$$

(vi) $f^2 \in RI[a,b]$;

(vii) $fg \in RI[a,b]$;

(viii) *If there exists* m, M *such that* $0 < m \leq\mid f \mid\leq M$, *then* $1/f \in RI[a,b]$.

Proof.
(i)

Subproof. *Note given a partition* $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$, *for each* $j = 1, \ldots, p$ *we can easily show that the supremum over order pairs can be computed in either order.*

$$\begin{aligned} \sup_{x,y \in [x_{j-1}, x_j]} (f(x) - f(y)) &= \sup_{y \in [x_{j-1}, x_j]} \sup_{y \in [x_{j-1}, x_j]} (f(x) - f(y)) \\ &= \sup_{x \in [x_{j-1}, x_j]} \sup_{y \in [x_{j-1}, x_j]} (f(x) - f(y)) \end{aligned}$$

Thus,

$$\begin{aligned} \sup_{x,y \in [x_{j-1}, x_j]} (f(x) - f(y)) &= \sup_{y \in [x_{j-1}, x_j]} \sup_{x \in [x_{j-1}, x_j]} (f(x) - f(y)) \\ &= \sup_{y \in [x_{j-1}, x_j]} (M_j - f(y)) \\ &= M_j + \sup_{y \in [x_{j-1}, x_j]} (-f(y)) \\ &= M_j - \inf_{y \in [x_{j-1}, x_j]} (f(y)) \\ &= M_j - m_j \end{aligned}$$

Now, let m'_j *and* M'_j *be defined by*

$$m'_j = \inf_{[x_{j-1},x_j]} \mid f(x) \mid$$

$$M'_j = \sup_{[x_{j-1},x_j]} \mid f(x) \mid .$$

Then, arguing as we did earlier, we find

$$M'_j - m'_j = \sup_{x,y \in [x_{j-1},x_j]} \mid f(x) \mid - \mid f(y) \mid .$$

Claim*:* $\sup_{x,y} \mid f(x) - f(y) \mid = M_j - m_j$
To see this is true, note

$$\mid f(x) - f(y) \mid = \begin{cases} f(x) - f(y), & f(x) \geq f(y) \\ f(y) - f(x), & f(x) < f(y) \end{cases}$$

In either case, we have $\mid f(x) - f(y) \mid \leq M_j - m_j$ *for all* x, y*, implying that* $\sup_{x,y} \mid f(x) - f(y) \mid \leq M_j - m_j$.

To see the reverse inequality holds, we first note that if $M_j = m_j$*, we see the reverse inequality holds trivially as* $\sup_{x,y} \mid f(x) - f(y) \mid \geq 0 = M_j - m_j$*. Hence, we may assume without loss of generality that the gap* $M_j - m_j$ *is positive.*
Then, given $0 < \epsilon < (1/2(M_j - m - j)$*, there exist,* $s, t \in [x_{j-1}, x_j]$ *such that* $M_j - \epsilon/2 < f(s)$ *and* $m_j + \epsilon/2 > f(t)$*, so that* $f(s) - f(t) > M_j - m_j - \epsilon$*. by our choice of* ϵ*, these terms are positive and so we also have* $\mid f(s) - f(t) \mid > M_j - m_j - \epsilon$*. It follows that*

$$\sup_{x,y \in [x_{j-1},x_j]} \mid f(x) - f(y) \mid \geq \mid f(s_j) - f(t_j) \mid > M_j - m_j - \epsilon \mid .$$

Since we can make ϵ *arbitrarily small, this implies that*

$$\sup_{x,y \in [x_{j-1},x_j]} \mid f(x) - f(y) \mid \geq M_j - m_j.$$

This establishes the reverse inequality and proves the claim $\lozenge$*.*

Thus, for each $j = 1, \ldots, p$*, we have*

$$M_j - m_j = \sup_{x,y \in [x_{j-1},x_j]} \mid f(x) - f(y) \mid .$$

So, since $\mid f(x) \mid - \mid f(y) \mid \leq \mid f(x) - f(y) \mid$ *for all* x, y*, it follows that* $M'_j - m'_j \leq M_j - m_j$*, implying that* $\sum_{\boldsymbol{\pi}} (M'_j - m'_j) \Delta x_j \leq \sum_{\boldsymbol{\pi}} (M_j - m_j) \Delta x_j$*. Since* f *is integrable by hypothesis, by Theorem 4.2.4, we know the Riemann criterion must also hold for* $\mid f \mid$*. Hence,* $\mid f \mid$ *is Riemann integrable.* □

The other results now follow easily. (ii)

Subproof. *We have* $f \leq \mid f \mid$ *and* $f \geq - \mid f \mid$*, so that*

$$\int_a^b f(x)dx \leq \int_a^b \mid f(x) \mid dx$$
$$\int_a^b f(x)dx \geq -\int_a^b \mid f(x) \mid dx,$$

from which it follows that

$$-\int_a^b \mid f(x) \mid dx \leq \int_a^b f(x)dx \leq \int_a^b \mid f(x) \mid dx$$

and so

$$\left|\int_a^b f\right| \leq \int_a^b \mid f \mid,$$

□

(iii) and (iv)

Subproof. *This follows from the facts that* $f^+ = \frac{1}{2}(\mid f \mid + f)$ *and* $f^- = \frac{1}{2}(\mid f \mid - f)$ *and the Riemann integral is a linear mapping.* □

(v)

Subproof. *This follows from the facts that* $f = f^+ - f^-$ *and* $\mid f \mid = f^+ + f^-$ *and the linearity of the integral.* □

(vi)

Subproof. *Note that, since* f *is bounded, there exists* $K > 0$ *such that* $\mid f(x) \mid \leq K$ *for all* $x \in [a, b]$. *Consequently, for all* $x, y \in [a, b]$, *we have* $\mid (f(x))^2 - (f(y))^2 \mid \leq 2K \mid f(x) - f(y) \mid$. *Thus, the integrability of* f *and the Riemann criterion imply that* f^2 *is integrable.* □

(vii)

Subproof. *To prove that* fg *is integrable when* f *and* g *are, simply note that*

$$fg = (1/2)\bigg((f+g)^2 - f^2 - g^2\bigg).$$

Property (vi) and the linearity of the integral then imply fg *is integrable.* □

(viii)

Subproof. *Suppose* $f \in RI[a, b]$ *and there exist* $M, m > 0$ *such that* $m \leq \mid f(x) \mid \leq M$ *for all* $x \in [a, b]$. *Note that*

$$\frac{1}{f(x)} - \frac{1}{f(y)} = \frac{f(y) - f(x)}{f(x)f(y)}.$$

Let $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$ *be a partition of* $[a, b]$, *and define*

$$\begin{aligned} M'_j &= \sup_{[x_{j-1},x_j]} \frac{1}{f(x)} \\ m'_j &= \inf_{[x_{j-1},x_j]} \frac{1}{f(x)}. \end{aligned}$$

Then we have

$$\begin{aligned} M'_j - m'_j &= \sup_{x,y \in [x_{j-1},x_j]} \frac{f(y)-f(x)}{f(x)f(y)} \\ &\leq \sup_{x,y \in [x_{j-1},x_j]} \frac{\mid f(y)-f(x) \mid}{\mid f(x) \mid\mid f(y) \mid} \\ &\leq \frac{1}{m^2} \sup_{x,y \in [x_{j-1},x_j]} \mid f(y)-f(x) \mid \\ &\leq \frac{M_j - m_j}{m^2}. \end{aligned}$$

Since $f \in RI[a,b]$*, given* $\epsilon > 0$ *there is a partition* $\boldsymbol{\pi}_0$ *such that* $U(f,\boldsymbol{\pi}) - L(f,\boldsymbol{\pi}) < m^2\epsilon$ *for any refinement, pi, of* $\boldsymbol{\pi}_0$*. Hence, the previous inequality implies that, for any such refinement, we have*

$$\begin{aligned} U\Big(\frac{1}{f},\boldsymbol{\pi}\Big) - L\Big(\frac{1}{f},\boldsymbol{\pi}\Big) &= \sum_{\boldsymbol{\pi}} (M'_j - m'_j)\Delta x_j \\ &\leq \frac{1}{m^2} \sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j \\ &\leq \frac{1}{m^2}\Big(U(f,\boldsymbol{\pi}) - L(f,\boldsymbol{\pi})\Big) \\ &< \frac{m^2\epsilon}{m^2} = \epsilon. \end{aligned}$$

Thus $1/f$ *satisfies the Riemann Criterion and hence it is integrable.* □

■

4.4 What Functions Are Riemann Integrable?

Now we need to show that the set $RI[a,b]$ is nonempty. We begin by showing that all continuous functions on $[a,b]$ will be Riemann Integrable.

Theorem 4.4.1. Continuous Implies Riemann Integrable

If $f \in C[a,b]$*, then* $f \in RI[a,b]$*.*

Proof. *Since* f *is continuous on a compact set, it is uniformly continuous. Hence, given* $\epsilon > 0$*, there is a* $\delta > 0$ *such that* $x, y \in [a,b]$*,* $\mid x - y \mid < \delta \Rightarrow \mid f(x) - f(y) \mid < \epsilon/(b-a)$*. Let* $\boldsymbol{\pi}_0$ *be a partition such that* $\mid\mid \boldsymbol{\pi}_0 \mid\mid < \delta$*, and let* $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$ *be any refinement of* $\boldsymbol{\pi}_0$*. Then* $\boldsymbol{\pi}$ *also satisfies* $\mid\mid \boldsymbol{\pi} \mid\mid < \delta$*.*

Since f is continuous on each subinterval $[x_{j-1}, x_j]$, f attains its supremum, M_j, and infimum, m_j, at points s_j and t_j, respectively. That is, $f(s_j) = M_j$ and $f(t_j) = m_j$ for each $j = 1, \ldots, p$. Thus, the uniform continuity of f on each subinterval implies that, for each j,

$$M_j - m_j = \mid f(s_j) - f(t_j) \mid < \frac{\epsilon}{b-a}.$$

Thus, we have

$$U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) = \sum_{\boldsymbol{\pi}} (M_j - m_j) \Delta x_j < \frac{\epsilon}{b-a} \sum_{\boldsymbol{\pi}} \Delta x_j = \epsilon.$$

Since $\boldsymbol{\pi}$ was an arbitrary refinement of $\boldsymbol{\pi}_0$, it follows that f satisfies Riemann's criterion. Hence, $f \in RI[a, b]$. ■

Theorem 4.4.2. Constant Functions Are Riemann Integrable

If $f : [a, b] \rightarrow \Re$ is a constant function, $f(t) = c$ for all t in $[a, b]$, then f is Riemann Integrable on $[a, b]$ and $\int_a^b f(t)dt = c(b - a)$.

Proof. *For any partition $\boldsymbol{\pi}$ of $[a, b]$, since f is a constant, all the individual m_j's and M_j's associated with $\boldsymbol{\pi}$ take on the value c. Hence, $U(f, \boldsymbol{\pi}) - U(f, \boldsymbol{\pi}) = 0$ always. It follows immediately that f satisfies the Riemann Criterion and hence is Riemann Integrable. Finally, since f is integrable, by Theorem 4.1.2, we have*

$$c(b - a) \leq RI(f; a, b) \leq c(b - a).$$

Thus, $\int_a^b f(t)dt = c(b - a)$. ■

Theorem 4.4.3. Monotone Implies Riemann Integrable

If f is monotone on $[a, b]$, then $f \in RI[a, b]$.

Proof. *As usual, for concreteness, we assume that f is monotone increasing. We also assume $f(b) > f(a)$, for if not, then f is constant and must be integrable by Theorem 4.4.2. Let $\epsilon > 0$ be given, and let $\boldsymbol{\pi}_0$ be a partition of $[a, b]$ such that $\| \boldsymbol{\pi}_0 \| < \epsilon / (f(b) - f(a))$. Let $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$ be any refinement of $\boldsymbol{\pi}_0$. Then $\boldsymbol{\pi}$ also satisfies $\| \boldsymbol{\pi} \| < \epsilon / (f(b) - f(a))$. Thus, for each $j = 1, \ldots, p$, we have*

$$\Delta x_j < \frac{\epsilon}{f(b) - f(a)}.$$

Since f is increasing, we also know that $M_j = f(x_j)$ and $m_j = f(x_{j-1})$ for each j. Hence,

$$\begin{aligned}
U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) &= \sum_{\boldsymbol{\pi}} (M_j - m_j) \Delta x_j \\
&= \sum_{\boldsymbol{\pi}} [f(x_j) - f(x_{j-1})] \Delta x_j \\
&< \frac{\epsilon}{f(b) - f(a)} \sum_{\boldsymbol{\pi}} [f(x_j) - f(x_{j-1})].
\end{aligned}$$

But this last sum is telescoping and sums to $f(b) - f(a)$*. So, we have*

$$U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) < \frac{\epsilon}{f(b) - f(a)}(f(b) - f(a)) = \epsilon.$$

Thus, f *satisfies Riemann's criterion.* ■

Theorem 4.4.4. Bounded Variation Implies Riemann Integrable

If $f \in BV[a, b]$*, then* $f \in RI[a, b]$*.*

Proof. *Since* f *is of bounded variation, there are functions* u *and* v*, defined on* $[a, b]$ *and both monotone increasing, such that* $f = u - v$*. Hence, by the linearity of the integral and the previous theorem,* $f \in RI[a, b]$*.* ■

4.5 Further Properties of the Riemann Integral

We first want to establish the familiar summation property of the Riemann integral over an interval $[a, b] = [a, c] \cup [c, b]$. Most of the technical work for this result is done in the following Lemma.

Lemma 4.5.1. The Upper And Lower Darboux Integral Is Additive On Intervals

Let $f \in B[a, b]$ *and let* $c \in (a, b)$*. Let*

$$\underline{\int_a^b} f(x)\, dx = L(f) \text{ and } \overline{\int_a^b} f(x)\, dx = U(f)$$

denote the lower and upper Darboux integrals of f *on* $[a, b]$*, respectively. Then we have*

$$\overline{\int_a^b} f(x)dx = \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx$$

$$\underline{\int_a^b} f(x)dx = \underline{\int_a^c} f(x)dx + \underline{\int_c^b} f(x)dx.$$

Proof. *We prove the result for the upper integrals as the lower integral case is similar. Let* $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a, b]$ *be given by* $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$*. We first assume that* c *is a partition point of* $\boldsymbol{\pi}$*. Thus, there is some index* $1 \leq k_0 \leq p-1$ *such that* $x_{k_0} = c$*. For any interval* $[\alpha, \beta]$*, let* $U_\alpha^\beta(f, \boldsymbol{\pi})$ *denote the upper sum of* f *for the partition* $\boldsymbol{\pi}$ *over* $[\alpha, \beta]$*. Now, we can rewrite* $\boldsymbol{\pi}$ *as* $\boldsymbol{\pi} = \{x_0, x_1, \ldots, x_{k_0}\} \cup \{x_{k_0}, x_{k_0+1}, \ldots, x_p\}$*. Let* $\boldsymbol{\pi}_1 = \{x_0, \ldots, x_{k_0}\}$ *and* $\boldsymbol{\pi}_2 = \{x_{k_0}, \ldots, x_p\}$*. Then* $\boldsymbol{\pi}_1 \in \boldsymbol{\Pi}[a, c]$*,* $\boldsymbol{\pi}_2 \in \boldsymbol{\Pi}[c, b]$*, and*

$$\begin{aligned} U_a^b(f, \boldsymbol{\pi}) &= U_a^c(f, \boldsymbol{\pi}_1) + U_c^b(f, \boldsymbol{\pi}_2) \\ &\geq \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx, \end{aligned}$$

by the definition of the upper sum. Now, if c *is not in* $\boldsymbol{\pi}$*, then we can refine* $\boldsymbol{\pi}$ *by adding* c*, obtaining the partition* $\boldsymbol{\pi}' = \{x_0, x_1, \ldots, x_{k_0}, c, x_{k_0+1}, \ldots, x_p\}$*. Splitting up* $\boldsymbol{\pi}'$ *at* c *as we did before into* $\boldsymbol{\pi}_1$ *and*

$\boldsymbol{\pi}_2$, *we see that* $\boldsymbol{\pi'} = \boldsymbol{\pi}_1 \vee \boldsymbol{\pi}_2$ *where* $\boldsymbol{\pi}_1 = \{x_0, \ldots, x_{k_0}, c\}$ *and* $\boldsymbol{\pi}_2 = \{c, x_{k_0+1}, \ldots, x_p\}$. *Thus, by our properties of upper sums, we see that*

$$U_a^b(f, \boldsymbol{\pi}) \geq U_a^b(f, \boldsymbol{\pi'}) = U_a^c(f, \boldsymbol{\pi}_1) + U_c^b(f, \boldsymbol{\pi}_2) \geq \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx.$$

Combining both cases, we can conclude that for any partition $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]$, *we have*

$$U_a^b(f, \boldsymbol{\pi}) \geq \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx,$$

which implies that

$$\overline{\int_a^b} f(x)dx \geq \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx.$$

Now we want to show the reverse inequality. Let $\epsilon > 0$ *be given. By the definition of the upper integral, there exists* $\boldsymbol{\pi}_1 \in \boldsymbol{\Pi}[a,c]$ *and* $\boldsymbol{\pi}_2 \in [c,b]$ *such that*

$$\begin{aligned} U_a^c(f, \boldsymbol{\pi}_1) &< \overline{\int_a^c} f(x)dx + \frac{\epsilon}{2} \\ U_c^b(f, \boldsymbol{\pi}_2) &< \overline{\int_c^b} f(x)dx + \frac{\epsilon}{2}. \end{aligned}$$

Let $\boldsymbol{\pi} = \boldsymbol{\pi}_1 \cup \boldsymbol{\pi}_2 \in \boldsymbol{\Pi}[a,b]$. *It follows that*

$$U_a^b(f, \boldsymbol{\pi}) = U_a^c(f, \boldsymbol{\pi}_1) + U_c^b(f, \boldsymbol{\pi}_2) < \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx + \epsilon.$$

But, by definition, we have

$$\overline{\int_a^b} f(x)dx \leq U_a^b(f, \boldsymbol{\pi})$$

for all $\boldsymbol{\pi}$. *Hence, we see that*

$$\overline{\int_a^b} f(x)dx < \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx + \epsilon.$$

Since ϵ *was arbitrary, this proves the reverse inequality we wanted. We can conclude, then, that*

$$\overline{\int_a^b} f(x)dx = \overline{\int_a^c} f(x)dx + \overline{\int_c^b} f(x)dx.$$

■

Theorem 4.5.2. The Riemann Integral Exists On Subintervals

If $f \in RI[a,b]$ *and* $c \in (a,b)$, *then* $f \in RI[a,c]$ *and* $f \in RI[c,b]$.

Proof. *Let $\epsilon > 0$ be given. Then there is a partition $\boldsymbol{\pi}_0 \in \boldsymbol{\Pi}[a,b]$ such that $U_a^b(f,\boldsymbol{\pi}) - L_a^b(f,\boldsymbol{\pi}) < \epsilon$ for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$. Let $\boldsymbol{\pi}_0$ be given by $\boldsymbol{\pi}_0 = \{x_0 = a, x_1, \ldots, x_p = b\}$. Define $\boldsymbol{\pi}_0' = \boldsymbol{\pi}_0 \cup \{c\}$, so there is some index k_0 such that $x_{k_0} \leq c \leq x_{k_0+1}$. Let $\boldsymbol{\pi}_1 = \{x_0, \ldots, x_{k_0}, c\}$ and $\boldsymbol{\pi}_2 = \{c, x_{k_0+1}, \ldots, x_p\}$. Then $\boldsymbol{\pi}_1 \in \boldsymbol{\Pi}[a,c]$ and $\boldsymbol{\pi}_2 \in \boldsymbol{\Pi}[c,b]$. Let $\boldsymbol{\pi}_1'$ be a refinement of $\boldsymbol{\pi}_1$. Then $\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2$ is a refinement of $\boldsymbol{\pi}_0$, and it follows that*

$$\begin{aligned} U_a^c(f,\boldsymbol{\pi}_1') - L_a^c(f,\boldsymbol{\pi}_1') &= \sum_{\boldsymbol{\pi}_1'} (M_j - m_j)\Delta x_j \\ &\leq \sum_{\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2} (M_j - m_j)\Delta x_j \\ &\leq U_a^b(f,\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2) - L_a^b(f,\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2). \end{aligned}$$

But, since $\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2$ refines $\boldsymbol{\pi}_0$, we have

$$U_a^b(f,\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2) - L_a^b(f,\boldsymbol{\pi}_1' \cup \boldsymbol{\pi}_2) < \epsilon,$$

implying that

$$U_a^c(f,\boldsymbol{\pi}_1') - L_a^c(f,\boldsymbol{\pi}_1') < \epsilon$$

for all refinements, $\boldsymbol{\pi}_1'$, of $\boldsymbol{\pi}_1$. Thus, f satisfies Riemann's criterion on $[a,c]$, and $f \in RI[a,c]$. The proof on $[c,b]$ is done in exactly the same way. ■

Theorem 4.5.3. The Riemann Integral Is Additive On Subintervals

If $f \in RI[a,b]$ and $c \in (a,b)$, then

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Proof. *Since $f \in RI[a,b]$, we know that*

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

Further, we also know that $f \in RI[a,c]$ and $f \in RI[c,b]$ for any $c \in (a,b)$. Thus,

$$\begin{aligned} \overline{\int_a^c} f(x)dx &= \underline{\int_a^c} f(x)dx \\ \overline{\int_c^b} f(x)dx &= \underline{\int_c^b} f(x)dx. \end{aligned}$$

So, applying Lemma 4.5.1, we conclude that, for any $c \in (a,b)$,

$$\int_a^b f(x)dx = \overline{\int_a^b} f(x)\,dx = \overline{\int_a^c} f(x)\,dx + \overline{\int_c^b} f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

■

4.6 The Fundamental Theorem Of Calculus

The next result is the well-known **Fundamental of Theorem of Calculus**.

Theorem 4.6.1. The Fundamental Theorem Of Calculus
Let $f \in RI[a,b]$. Define $F : [a,b] \to \Re$ by

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

Then

(i) $F \in BV[a,b]$;

(ii) $F \in C[a,b]$;

(iii) if f is continuous at $c \in [a,b]$, then F is differentiable at c and $F'(c) = f(c)$.

Proof. *First, note that $f \in RI[a,b] \Rightarrow f \in R[a,x]$ for all $x \in [a,b]$, by our previous results. Hence, F is well-defined. We will prove the results in order.* (*i*)

Subproof. *Let $\boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]$ be given by $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$. Then the fact that $f \in R[a,x_j]$ implies that $f \in R[x_{j-1}, x_j]$ for each $j = 1, \ldots, p$. Thus, we have*

$$m_j \Delta x_j \leq \int_{x_{j-1}}^{x_j} f(t)dt \leq M_j \Delta x_j.$$

This implies that, for each j, we have

$$\left| \int_{x_{j-1}}^{x_j} f(t)dt \right| \leq \| f \|_\infty \Delta x_j.$$

Thus,

$$\begin{aligned} \mid \Delta F_j \mid &= \mid F(x_j) - F(x_{j-1}) \mid \\ &= \left| \int_a^{x_j} f(t)dt - \int_a^{x_{j-1}} f(t)dt \right| \\ &= \left| \int_{x_{j-1}}^{x_j} f(t)dt \right| \\ &\leq \| f \|_\infty \Delta x_j. \end{aligned}$$

Summing over $\boldsymbol{\pi}$, we obtain

$$\sum_{\boldsymbol{\pi}} \mid \Delta F_j \mid \leq \| f \|_\infty \sum_{\boldsymbol{\pi}} \Delta x_j = (b-a) \| f \|_\infty < \infty.$$

Since the partition $\boldsymbol{\pi}$ was arbitrary, we conclude that $F \in BV[a,b]$. □

(ii)

Subproof. *Now, let $x, y \in [a,b]$ be such that $x < y$. Then*

$$\inf_{[x,y]} f(t)\,(y-x) \leq \int_x^y f(t)dt \leq \sup_{[x,y]} f(t)\,(y-x),$$

which implies that

$$| F(y) - F(x) | = \left| \int_x^y f(t)dt \right| \leq || f ||_\infty \ (y-x).$$

A similar argument shows that if $y, x \in [a,b]$ satisfy $y < x$, then

$$| F(y) - F(x) | = \left| \int_x^y f(t)dt \right| \leq || f ||_\infty \ (x-y).$$

Let $\epsilon > 0$ be given. Then if

$$| x - y | < \frac{\epsilon}{|| f ||_\infty + 1},$$

we have

$$| F(y) - F(x) | \leq || f ||_\infty | y - x | < \frac{|| f ||_\infty}{|| f ||_\infty + 1} \epsilon < \epsilon.$$

Thus, F is continuous at x and, consequently, on $[a,b]$. □

(iii)

Subproof. *Finally, assume f is continuous at $c \in [a,b]$, and let $\epsilon > 0$ be given. Then there exists $\delta > 0$ such that $x \in (c-\delta, c+\delta) \cap [a,b]$ implies $| f(x) - f(c) | < \epsilon/2$. Pick $h \in \Re$ such that $0 < | h | < \delta$ and $c + h \in [a,b]$. Let's assume, for concreteness, that $h > 0$. Define*

$$m = \inf_{[c,c+h]} f(t) \qquad \textit{and} \qquad M = \sup_{[c,c+h]} f(t).$$

If $c < x < c+h$, then we have $x \in (c-\delta, c+\delta) \cap [a,b]$ and $-\epsilon/2 < f(x) - f(c) < \epsilon/2$. That is,

$$f(c) - \frac{\epsilon}{2} < f(x) < f(c) + \frac{\epsilon}{2} \qquad \forall x \in [c, c+h].$$

Hence, $m \geq f(c) - \epsilon/2$ and $M \leq f(c) + \epsilon/2$. Now, we also know that

$$mh \leq \int_c^{c+h} f(t)dt \leq Mh.$$

Thus, we have

$$\frac{F(c+h) - F(c)}{h} = \frac{\int_a^{c+h} f(t)dt - \int_a^c f(t)dt}{h} = \frac{\int_c^{c+h} f(t)dt}{h}.$$

Combining inequalities, we find

$$f(c) - \frac{\epsilon}{2} \leq m \leq \frac{F(c+h) - F(c)}{h} \leq M \leq f(c) + \frac{\epsilon}{2}$$

yielding

$$\Rightarrow \left| \frac{F(c+h)-F(c)}{h} - f(c) \right| \le \frac{\epsilon}{2} < \epsilon$$

if $x \in [c, c+h]$.
The case where $h < 0$ *is handled in exactly the same way. Thus, since* ϵ *was arbitrary, this shows that* F *is differentiable at* c *and* $F'(c) = f(c)$. *Note that if* $c = a$ *or* $c = b$, *we need only consider the definition of the derivative from one side.* □

■

Comment 4.6.1. *We call* $F(x)$ *the indefinite integral of* f. F *is always better behaved than* f, *since integration is a smoothing operation. We can see that* f *need not be continuous, but, as long as it is integrable,* F *is always continuous.*

The next result is one of the many mean value theorems in the theory of integration. It is a more general form of the standard mean value theorem given in beginning calculus classes.

Theorem 4.6.2. The Mean Value Theorem For Riemann Integrals
Let $f \in C[a, b]$, *and let* $g \ge 0$ *be integrable on* $[a, b]$. *Then there is a point,* $c \in [a, b]$, *such that*

$$\int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx.$$

Proof. *Since* f *is continuous, it is also integrable. Hence,* fg *is integrable. Let* m *and* M *denote the lower and upper bounds of* f *on* $[a, b]$, *respectively. Then* $mg(x) \le f(x)g(x) \le Mg(x)$ *for all* $x \in [a, b]$. *Since the integral preserves order, we have*

$$m \int_a^b g(x)dx \le \int_a^b f(x)g(x)dx \le M \int_a^b g(x)dx.$$

If the integral of g *on* $[a, b]$ *is* 0, *then this shows that the integral of* fg *will also be* 0. *Hence, in this case, we can choose any* $c \in [a, b]$ *and the desired result will follow. If the integral of* g *is not* 0, *then it must be positive, since* $g \ge 0$. *Hence, we have, in this case,*

$$m \le \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \le M.$$

Now, f *must be uniformly continuous, implying that it attains the values* M *and* m *at some points. Hence, by the intermediate value theorem, there must be some* $c \in [a, b]$ *such that*

$$f(c) = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx}.$$

This implies the desired result. ■

The next result is another standard mean value theorem from basic calculus. It is a direct consequence of the previous theorem, by simply letting $g(x) = 1$ for all $x \in [a, b]$. This result can be interpreted as stating that integration is an averaging process.

Theorem 4.6.3. Average Value For Riemann Integrals

If $f \in C[a,b]$, then there is a point $c \in [a,b]$ such that

$$\frac{1}{b-a}\int_a^b f(x)dx = f(c).$$

The next result is the standard means for calculating definite integrals in basic calculus. We start with a definition.

Definition 4.6.1. The Antiderivative of f

Let $f : [a,b] \to \Re$ be a bounded function. Let $G : [a,b] \to \Re$ be such that G' exists on $[a,b]$ and $G'(x) = f(x)$ for all $x \in [a,b]$. Such a function is called an **antiderivative** *or a* **primitive** *of f.*

Comment 4.6.2. *The idea of an antiderivative is intellectually distinct from the Riemann integral of a bounded function f. Consider the following function f defined on $[-1,1]$.*

$$f(x) = \begin{cases} x^2\sin(1/x^2), & x \neq 0,\ x \in [-1,1] \\ 0, & x = 0 \end{cases}$$

It is easy to see that this function has a removable discontinuity at 0. Moreover, f is even differentiable on $[-1,1]$ with derivative

$$f'(x) = \begin{cases} 2x\sin(1/x^2) - (2/x)\cos(1/x^2), & x \neq 0,\ x \in [-1,1] \\ 0, & x = 0 \end{cases}$$

Note f' is not *bounded on $[-1,1]$ and hence it can not be Riemann Integrable. Now to connect this to the idea of antiderivatives, just relabel the functions. Let g be defined by*

$$g(x) = \begin{cases} 2x\sin(1/x^2) - (2/x)\cos(1/x^2), & x \neq 0,\ x \in [-1,1] \\ 0, & x = 0 \end{cases}$$

then define G by

$$G(x) = \begin{cases} x^2\sin(1/x^2), & x \neq 0,\ x \in [-1,1] \\ 0, & x = 0 \end{cases}$$

We see that G is the antiderivative of g even though g itself does not have a Riemann integral. Again, the point is that the idea of the antiderivative of a function is intellectually distinct from that of being Riemann integrable.

Theorem 4.6.4. Cauchy's Fundamental Theorem

Let $f : [a,b] \to \Re$ be integrable. Let $G : [a,b] \to \Re$ be any antiderivative of f. Then

$$\int_a^b f(t)dt = G(t)\Big|_a^b = G(b) - G(a).$$

Proof. *Since G' exists on $[a, b]$, G must be continuous on $[a, b]$. Let $\epsilon > 0$ be given. Since f is integrable, there is a partition $\boldsymbol{\pi}_0 \in \mathbf{\Pi}[a, b]$ such that for any refinement, $\boldsymbol{\pi}$, of $\boldsymbol{\pi}_0$ and any $\boldsymbol{\sigma} \subset \boldsymbol{\pi}$, we have*

$$\left| S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) - \int_a^b f(x)dx \right| < \epsilon.$$

Let $\boldsymbol{\pi}$ be any refinement of $\boldsymbol{\pi}_0$, given by $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p = b\}$. The Mean Value Theorem for differentiable functions then tells us that there is an $s_j \in (x_{j-1}, x_j)$ such that $G(x_j) - G(x_{j-1}) = G'(s_j)\Delta x_j$. Since $G' = f$, we have $G(x_j) - G(x_{j-1}) = f(s_j)\Delta x_j$ for each $j = 1, \ldots, p$. The set of points, $\{s_1, \ldots, s_p\}$, is thus an evaluation set associated with $\boldsymbol{\pi}$. Hence,

$$\sum_{\boldsymbol{\pi}} [G(x_j) - G(x_{j-1})] = \sum_{\boldsymbol{\pi}} G'(s_j)\Delta x_j = \sum_{\boldsymbol{\pi}} f(s_j)\Delta x_j$$

The first sum on the left is a collapsing sum, hence we have

$$\Rightarrow G(b) - G(a) = S(f, \boldsymbol{\pi}, \{s_1, \ldots, s_p\}).$$

We conclude

$$\left| G(b) - G(a) - \int_a^b f(x)dx \right| < \epsilon.$$

Since ϵ was arbitrary, this implies the desired result. ■

Comment 4.6.3. *Not all functions (in fact, most functions) will have closed form, or analytically obtainable, antiderivatives. So, the previous theorem will not work in such cases.*

Theorem 4.6.5. The Recapture Theorem

If f is differentiable on $[a, b]$, and if $f' \in RI[a, b]$, then

$$\int_a^x f'(t)dt = f(x) - f(a).$$

Proof. *f is an antiderivative of f. Now apply Cauchy's Fundamental Theorem 4.6.4.* ■

Another way to evaluate Riemann integrals is to directly approximate them using an appropriate sequence of partitions. Theorem 4.6.6 is a fundamental tool that tells us when and why such approximations will work.

Theorem 4.6.6. Approximation Of The Riemann Integral

If $f \in RI[a, b]$, then given any sequence of partitions $\{\boldsymbol{\pi_n}\}$ with any associated sequence of evaluation sets $\{\boldsymbol{\sigma_n}\}$ that satisfies $\|\boldsymbol{\pi_n}\| \to 0$, we have

$$\lim_{n \to \infty} S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) = \int_a^b f(x)\, dx$$

Proof. *Since f is integrable, given a positive ϵ, there is a partition $\boldsymbol{\pi_0}$ so that*

$$| S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}) - \int_a^b f(x)dx | \quad < \quad \epsilon/2, \ \boldsymbol{\pi_0} \preceq \boldsymbol{\pi}, \ \boldsymbol{\sigma} \subseteq \boldsymbol{\pi}. \qquad (*)$$

Let the partition $\boldsymbol{\pi_0}$ be $\{x_0, x_1, \ldots, x_P\}$ and let ξ be defined to be the smallest Δx_j from $\boldsymbol{\pi_0}$. Then since the norm of the partitions $\boldsymbol{\pi_n}$ goes to zero, there is a positive integer N so that

$$|| \boldsymbol{\pi_n} || \quad < \quad \min(\xi, \epsilon/(4P \,||\, f \,||)\infty)) \qquad (*)$$

Now pick any $n > N$ and label the points of $\boldsymbol{\pi_n}$ as $\{y_0, y_1, \ldots, y_Q\}$. We see that the points in $\boldsymbol{\pi_n}$ are close enough together so that at most one *point of $\boldsymbol{\pi_0}$ lies in any subinterval $y_{j-1}, y_j]$ from $\boldsymbol{\pi_n}$. This follows from our choice of ξ. So the intervals of $\boldsymbol{\pi_n}$ split into two pieces: those containing a point of $\boldsymbol{\pi_0}$ and those that do not have a $\boldsymbol{\pi_0}$ inside. Let $\mathscr{A}$ be the first collection of intervals and $\mathscr{B}$, the second. Note there are P points in $\boldsymbol{\pi_0}$ and so there are P subintervals in $\mathscr{B}$. Now consider the common refinement $\boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}$. The points in the common refinement match $\boldsymbol{\pi_n}$ except on the subintervals from $\mathscr{B}$. Let $[y_{j-1}, y_j]$ be such a subinterval and let γ_j denote the point from $\boldsymbol{\pi_0}$ which is in this subinterval. Let's define an evaluation set $\boldsymbol{\sigma}$ for this refinement $\boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}$ as follows.*

1. *if we are in the subintervals labeled $\mathscr{A}$, we choose as our evaluation point, the evaluation point s_j that is already in this subinterval since $\boldsymbol{\sigma_n} \subseteq \boldsymbol{\pi_n}$. Here, the length of the subinterval will be denoted by $\delta_j(\mathscr{A})$ which equals $y_j - y_{j-1}$ for appropriate indices.*

2. *if we are in the the subintervals labeled $\mathscr{B}$, we have two intervals to consider as $[y_{j-1}, y_j] = [y_{j-1}, \gamma_j] \cup [\gamma_j, y_j]$. Choose the evaluation point γ_j for both $[y_{j-1}, \gamma]$ and $[\gamma, y_j]$. Here, the length of the subintervals will be denoted by $\delta_j(\mathscr{B})$. Note that $\delta_j(\mathscr{B}) = \gamma_j - y_{j-1}$ or $y_j - \gamma_j$.*

Then we have

$$\begin{aligned}
S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) &= \sum_{\mathscr{A}} f(s_j)\delta_j(\mathscr{A}) + \sum_{\mathscr{B}} f(\gamma_j)\delta_j(\mathscr{A}) \\
&= \sum_{\mathscr{A}} f(s_j)(y_j - y_{j-1}) + \sum_{\mathscr{B}} (f(\gamma_j)(y_j - \gamma_j) + f(\gamma_j)(\gamma_j - y_{j-1})) \\
&= \sum_{\mathscr{A}} f(s_j)(y_j - y_{j-1}) + \sum_{\mathscr{B}} (f(\gamma_j)(y_j - y_{j-1}))
\end{aligned}$$

Thus, since the Riemann sums over $\boldsymbol{\pi_n}$ and $\boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}$ with these choices of evaluation sets match on $\mathscr{A}$, we have using Equation $*$ that

$$\begin{aligned}
| S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) - S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) | &= | \sum_{\mathscr{A}} (f(s_j) - f(\gamma_j)(y_j - y_{j-1}) | \\
&\leq \sum_{\mathscr{A}} (| f(s_j) | + | f(\gamma_j) |) (y_j - y_{j-1}) | \\
&\leq P \, 2 \, || f ||_\infty \, || \pi_n || \\
&< P \, 2 \, || f ||_\infty \, \frac{\epsilon}{4P \, || f ||_\infty} \\
&= \epsilon/2
\end{aligned}$$

We conclude that for our special evaluation set $\boldsymbol{\sigma}$ *for the refinement* $\boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}$ *that*

$$\begin{aligned} \| \, S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) - \int_a^b f(x)dx \, | &= | \, S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) - S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) + S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) - \int_a^b f(x)dx \, | \\ &\leq | \, S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) - S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) \, | + | \, S(f, \boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}, \boldsymbol{\sigma}) - \int_a^b f(x)dx \, | \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

using Equation $*$ *as* $\boldsymbol{\pi_n} \vee \boldsymbol{\pi_0}$ *refines* $\boldsymbol{\pi_0}$*. Since we can do this analysis for any* $n > N$*, we see we have shown the desired result.* ■

4.6.1 Homework

Exercise 4.6.1. *Let* $f(x) = x^2$ *on the interval* $[-1, 3]$*. Use Theorem 4.6.6 to prove that* $\int_{-1}^3 f(x)dx = 28/3$.

Hint. *We know* f *is Riemann integrable because it is continuous and so this theorem can be applied. Use the uniform approximations* $x_i = -1 + 4i/n$ *for* $i = 0$ *to* $i = n$ *to define partitions* $\boldsymbol{\pi_n}$*. Then using left or right hand endpoints on each subinterval to define the evaluation set* $\boldsymbol{\sigma_n}$*, you can prove directly that* $\int_{-1}^3 x^2 dx = \lim \, S(f, \boldsymbol{\pi_n}, \boldsymbol{\sigma_n}) = 28/3$*. Make sure you tell me all the reasoning involved.* □

Exercise 4.6.2. *If* f *is continuous, evaluate*

$$\lim_{x \to a} \frac{x}{x-a} \int_a^x f(t)dt$$

Exercise 4.6.3. *Prove if* f *is continuous on* $[a, b]$ *and* $\int_a^b f(x)g(x)dx = 0$ *for all choices of integrable* g*, then* f *is identically* 0.

4.7 Substitution Type Results

Using the Fundamental Theorem of Calculus, we can derive many useful tools.

Theorem 4.7.1. Integration By Parts

Assume $u : [a, b] \to \Re$ *and* $v : [a, b] \to \Re$ *are differentiable on* $[a, b]$ *and* u' *and* v' *are integrable. Then*

$$\int_a^x u(t)v'(t) \, dt = u(t)v(t)\Big|_a^x - \int_a^x v(t)u'(t) \, dt$$

Proof. *Since* u *and* v *are differentiable on* $[a, b]$*, they are also continuous and hence, integrable. Now apply the product rule for differentiation to obtain*

$$(u(t)v(t))' = u'(t)v(t) + u(t)v'(t)$$

By Theorem 4.3.1, we know products of integrable functions are integrable. Also, the integral is linear. Hence, the integral of both sides of the equation above is defined. We obtain

$$\int_a^x (u(t)v(t))' \, dt \quad = \quad \int_a^x u'(t)v(t) \, dt \, + \int_a^x u(t)v'(t) \, dt$$

Since $(uv)'$ is integrable, we can apply the Recapture Theorem to see

$$u(t)v(t)\Big|_a^x \quad = \quad \int_a^x u'(t)v(t) \, dt \, + \int_a^x u(t)v'(t) \, dt$$

This is the desired result. ■

Theorem 4.7.2. Substitution In Riemann Integration
Let f be continuous on $[c,d]$ and u be continuously differentiable on $[a,b]$ with $u(a) = c$ and $u(b) = d$. Then

$$\int_c^d f(u) \, du \quad = \quad \int_a^b f(u(t)) \, u'(t) \, dt$$

Proof. *Let F be defined on $[c,d]$ by $F(u) = \int_c^u f(t)dt$. Then since f is continuous, F is continuous and differentiable on $[c,d]$ by the Fundamental Theorem of Calculus. We know $F'(u) = f(u)$ and so*

$$F'(u(t)) = f(u(t)), \; a \, \leq \, t \, \leq \, b$$

implying

$$F'(u(t)) \, u'(t) = f(u(t))u'(t) \, , \; a \, \leq \, t \, \leq \, b$$

By the Chain Rule for differentiation, we also know

$$(F \circ u)'(t) = F(u(t))u'(t) \, , \; a \, \leq \, t \, \leq \, b.$$

and hence $(F \circ u)'(t) = f(u(t))u'(t)$ on $[a,b]$.
Now define g on $[a,b]$ by

$$\begin{aligned} g(t) \quad &= \quad (f \circ u)(t) \, u'(t) \; = \; f(u(t)) \, u'(t) \\ &= \quad (F \circ u)'(t). \end{aligned}$$

Since g is continuous, g is integrable on $[a,b]$. Now define G on $[a,b]$ by $G(t) = (F \circ u)(t)$. Then $G'(t) = f(u(t))u'(t) = g(t)$ on $[a,b]$ and G' is integrable. Now, apply the Cauchy Fundamental Theorem of Calculus to G to find

$$\int_a^b g(t) \, dt \quad = \quad G(b) \, - \, G(a)$$

or

$$\begin{aligned}
\int_a^b f(u(t))\, u'(t)\, dt &= F(u(b)) - F(u(a)) \\
&= \int_c^{u(b)=d} f(t)dt - \int_c^{u(a)=c} f(t)dt \\
&= \int_c^d f(t)dt.
\end{aligned}$$

■

Theorem 4.7.3. Leibnitz's Rule

Let f be continuous on $[a,b]$, $u : [c,d] \to [a,b]$ be differentiable on $[c,d]$ and $v : [c,d] \to [a,b]$ be differentiable on $[c,d]$. Then

$$\left(\int_{u(x)}^{v(x)} f(t)\, dt \right)' = f(v(x))v'(x) - f(u(x)u'(x)$$

Proof. *Let F be defined on $[a,b]$ by $F(y) = \int_a^y f(t)dt$. Since f is continuous, F is also continuous and moreover, F is differentiable with $F'(y)) = f(y)$. Since v is differentiable on $[c,d]$, we can use the Chain Rule to find*

$$\begin{aligned}
(F \circ v)'(x) &= F'(v(x))\, v'(x) \\
&= f(v(x))\, v'(x)
\end{aligned}$$

This says

$$\left(\int_a^{v(x)} f(t)\, dt \right)' = f(v(x))v'(x)$$

Next, define G on $[a,b]$ by $G(y) = \int_y^b f(t)dt = \int_a^b f(t) - \int_a^y f(t)dt$. Apply the Fundamental Theorem of Calculus to conclude

$$G'(y) = -\left(\int_a^y f(t)dt \right) = -f(y)$$

Again, apply the Chain Rule to see

$$\begin{aligned}
(G \circ u)'(x) &= G'(u(x))\, u'(x) \\
&= -f(u(x))\, u'(x).
\end{aligned}$$

We conclude

$$\left(\int_{u(x)}^b f(t)\, dt \right)' = -f(u(x))u'(x)$$

Now combine these results as follows:

$$\int_a^b f(t)dt \quad = \quad \int_a^{v(x)} f(t)dt + \int_{v(x)}^{u(x)} f(t)dt + \int_{u(x)}^b f(t)dt$$

or

$$\begin{aligned}(F \circ v)(x) + (G \circ u)(x) - \int_a^b f(t)dt &= -\int_{v(x)}^{u(x)} f(t)dt \\ &= \int_{u(x)}^{v(x)} f(t)dt\end{aligned}$$

Then, differentiate both sides to obtain

$$\begin{aligned}(F \circ v)'(x) + (G \circ u)'(x) &= f(v(x))v'(x) \; - \; f(u(x))u'(x) \\ &= \left(\int_{u(x)}^{v(x)} f(t)dt \right)'\end{aligned}$$

which is the desired result. ■

4.8 When Do Two Functions Have The Same Integral?

The last results in this chapter seek to find conditions under which the integrals of two functions, f and g, are equal.

Lemma 4.8.1. f Zero On (a, b) Implies Zero Riemann Integral
Let $f \in B[a, b]$, with $f(x) = 0$ on (a, b). Then f is integrable on $[a, b]$ and

$$\int_a^b f(x)dx = 0.$$

Proof. *If f is identically 0, then the result is follows easily. Now, assume $f(a) \neq 0$ and $f(x)$ on $(a, b]$. Let $\epsilon > 0$ be given, and let $\delta > 0$ satisfy*

$$\delta < \frac{\epsilon}{\mid f(a) \mid}.$$

Let $\boldsymbol{\pi}_0 \in \mathbf{\Pi}[a, b]$ be any partition such that $\| \boldsymbol{\pi}_0 \| < \delta$. Let $\boldsymbol{\pi} = \{x_0 = a, x_1, \ldots, x_p\}$ be any refinement of $\boldsymbol{\pi}_0$. Then $U(f, \boldsymbol{\pi}) = \max(f(a), 0)\Delta x_1$ and $L(f, \boldsymbol{\pi}) = \min(f(a), 0)\Delta x_1$. Hence, we have

$$U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) = [\max(f(a), 0) - \min(f(a), 0)]\Delta x_1 = \mid f(a) \mid \Delta x_1.$$

But

$$\mid f(a) \mid \Delta x_1 < \mid f(a) \mid \delta < \mid f(a) \mid \frac{\epsilon}{\mid f(a) \mid} = \epsilon.$$

Hence, if $\boldsymbol{\pi}$ is any refinement of $\boldsymbol{\pi}_0$, we have $U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) < \epsilon$. This shows that $f \in RI[a, b]$. Further, we have

$$U(f,\boldsymbol{\pi}) = \max(f(a),0)\Delta x_1 \Rightarrow U(f) = \inf_{\boldsymbol{\pi}} U(f,\boldsymbol{\pi}) = 0,$$

since we can make Δx_1 *as small as we wish. Likewise, we also see that* $L(f) = \sup_{\boldsymbol{\pi}} L(f,\boldsymbol{\pi}) = 0$, *implying that*

$$U(f) = L(f) = \int_a^b f(x)dx = 0.$$

The case where $f(b) \neq 0$ *and* $f(x) = 0$ *on* $[a,b)$ *is handled in the same way. So, assume that* $f(a), f(b) \neq 0$ *and* $f(x) = 0$ *for* $x \in (a,b)$. *Let* $\epsilon > 0$ *be given, and choose* $\delta > 0$ *such that*

$$\delta < \frac{\epsilon}{2\max\{\mid f(a) \mid, \mid f(b) \mid\}}.$$

Let $\boldsymbol{\pi}_0$ *be a partition of* $[a,b]$ *such that* $\mid \boldsymbol{\pi}_0 \mid < \delta$, *and let* $\boldsymbol{\pi}$ *be any refinement of* $\boldsymbol{\pi}_0$. *Then*

$$U(f,\boldsymbol{\pi}) = \max(f(a),0)\Delta x_1 + \max(f(b),0)\Delta x_p$$

$$L(f,\boldsymbol{\pi}) = \min(f(a),0)\Delta x_1 + \min(f(b),0)\Delta x_p.$$

It follows that

$$\begin{aligned} U(f,\boldsymbol{\pi}) - L(f,pi) &= [\max(f(a),0) - \min(f(a),0)]\Delta x_1 + [\max(f(b),0) - \min(f(a),0)]\Delta x_p \\ &= \mid f(a) \mid \Delta x_1 + \mid f(b) \mid \Delta x_p \\ &< \mid f(a) \mid \delta + \mid f(b) \mid \delta \\ &< \epsilon. \end{aligned}$$

Since we can make Δx_1 *and* Δx_p *as small as we wish, we see*

$$\int_a^b f(x)dx = 0.$$

■

Lemma 4.8.2. $f = g$ **on** (a,b) **Implies Riemann Integrals Match**

Let $f, g \in RI[a,b]$ *with* $f(x) = g(x)$ *on* (a,b). *Then*

$$\int_a^b f(x)dx = \int_a^b g(x)dx.$$

Proof. *Let* $h = f - g$, *and apply the previous lemma.* ■

Theorem 4.8.3. Two Riemann Integrable Functions Match At All But Finitely Many Points Implies Integrals Match

Let $f, g \in RI[a,b]$, and assume that $f = g$ except at finitely many points $c_1, \ldots, c_k$. Then

$$\int_a^b f(x)dx = \int_a^b g(x)dx.$$

Proof. *We may re-index the points $\{c_1, \ldots, c_k\}$, if necessary, so that $c_1 < c_2 < \cdots < c_k$. Then apply Lemma 4.8.2 on the intervals (c_{j-1}, c_j) for all allowable j. This shows*

$$\int_{c_{j-1}}^{c_j} f(t)dt = \int_{c_{j-1}}^{c_j} g(t)dt.$$

Then, since

$$\int_a^b f(t)dt = \sum_{j=1}^{k} \int_{c_{j-1}}^{c_j} f(t)dt$$

the results follows. ■

Theorem 4.8.4. f Bounded and Continuous At All But One Point Implies f is Riemann Integrable

if f is bounded on $[a,b]$ and continuous except at one point c in $[a,b]$, then f is Riemann integrable.

Proof. *For convenience, we will assume that c is an interior point, i.e. c is in (a,b). We will show that f satisfies the Riemann Criterion and so it is Riemann integrable. Let $\epsilon > 0$ be given. Since f is bounded on $[a,b]$, there is a real number M so that $f(x) < M$ for all x in $[a,b]$. We know f is continuous on $[a, c-\epsilon/(6M)]$ and f is continuous on $[c+\epsilon/(6M), b]$. Thus, f is integrable on both of these intervals and f satisfies the Riemann Criterion on both intervals. For this ϵ there is a partition $\boldsymbol{\pi_0}$ of $[a, c-\epsilon/(6M)]$ so that*

$$U(f, \boldsymbol{P}) - L(f, \boldsymbol{P}) \quad < \quad \epsilon/3, \text{ if} \boldsymbol{\pi_0} \preceq \boldsymbol{P}$$

and there is a partition $\boldsymbol{\pi_1}$ of $[c+\epsilon/(6M), b]$ so that

$$U(f, \boldsymbol{Q}) - L(f, \boldsymbol{Q}) \quad < \quad \epsilon/3, \text{ if} \boldsymbol{\pi_0} \preceq \boldsymbol{Q}.$$

Let $\boldsymbol{\pi_2}$ be the partition we get by combining $\boldsymbol{\pi})_0$ with the points $\{c-\epsilon/(6M), c+\epsilon/(6M)\}$ and $\boldsymbol{\pi})_1$. Then, we see

$$\begin{aligned} U(f, \boldsymbol{\pi_2}) - L(f, \boldsymbol{\pi_2}) \quad &= \quad U(f, \boldsymbol{\pi_0}) - L(f, \boldsymbol{\pi_0}) \; + \left(\sup_{x \in [c-\epsilon/(6M), c+\epsilon/(6M)]} f(x) \right) \epsilon/3 \; + \; U(f, \boldsymbol{\pi_1}) - L(f, \boldsymbol{\pi_1}) \\ &< \quad \epsilon/3 + M\epsilon/(3M) + \epsilon/3 \; = \; \epsilon \end{aligned}$$

Then if $\boldsymbol{\pi_2} \preceq \boldsymbol{\pi}$ on $[a,b]$, we have

$$U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) \quad < \quad \epsilon$$

This shows f satisfies the Riemann criterion and hence is integrable if the discontinuity c is interior to $[a, b]$. The argument at $c = a$ and $c = b$ is similar but a bit simpler as it only needs to be done from one side. Hence, we conclude f is integrable on $[a, b]$ in all cases.. ∎

It is then easy to extend this result to a function f which is bounded and continuous on $[a, b]$ except at a finite number of points $\{x_1, x_2, \ldots, x_k\}$ for some positive integer k. We state this as Theorem 4.8.5.

Theorem 4.8.5. ***f* Bounded and Continuous At All But Finitely Many Points Implies *f* is Riemann Integrable**

if f is bounded on $[a, b]$ and continuous except at finitely many points $\{x_1, x_2, \ldots, x_k\}$ in $[a, b]$,
then f is Riemann integrable.

Proof. *We may assume without loss of generality that the points of discontinuity are ordered as $a < x_1 < x_2 < \ldots < x_k < b$. Then f is continuous except at x_1 on $[a, x_1]$ and hence by Theorem 4.8.4 f is integrable on $[a, x_1]$. Now apply this argument on each of the subintervals $x_{k-1}, x_k]$ in turn.* ∎

Chapter 5

Further Riemann Integration Results

In this chapter, we will explore certain aspects of Riemann Integration that are more subtle. We begin with a limit interchange theorem. A good reference for this is (Fulks (3) 1978) .

5.1 The Limit Interchange Theorem for Riemann Integration

Suppose you knew that the sequence of functions $\{x_n\}$ contained in $RI[a, b]$ converged uniformly to the function x on $[a, b]$. Is it true that $\int_a^b x(t)dt = \lim_{n\to\infty} x_n(t)dt$? The answer to this question is *Yes!* and it is our Theorem 5.1.1.

Theorem 5.1.1. The Riemann Integral Limit Interchange Theorem
Let $\{x_n\}$ be a sequence of Riemann Integrable functions on $[a, b]$ which converge uniformly to the function x on $[a, b]$. Then x is also Riemann Integrable on $[a, b]$ and

$$\int_a^b x(t)dt = \lim_{n\to\infty} x_n(t)dt$$

Proof. *First, we show that x is Riemann integrable on $[a, b]$. Let ϵ be given. Then since x_n converges uniformly to x on $[a, b]$,*

$$\exists\, \delta > 0 \ni |\, x_n(t) - x(t)\,| < \frac{\epsilon}{5(b-a)} \,\forall\, n > N,\, t \in [a, b] \qquad (\boldsymbol{\alpha})$$

Fix any $n_1 > N$. Then since x_{n_1} is integrable,

$$\exists\, \boldsymbol{\pi_0} \in \Pi[a, b] \ni U(x_{n_1}, \boldsymbol{\pi}) - L((x_{n_1}, \boldsymbol{\pi}) < frac\epsilon 5 \,\forall\, \boldsymbol{\pi_0} \preceq \boldsymbol{\pi} \qquad (\boldsymbol{\beta})$$

Since x_n converges uniformly to x on $[a,b]$, you should be able to show that x is bounded on $[a,b]$. Hence, we can define

$$\begin{aligned} M_j &= \sup_{[x_{j-1},x_j]} x(t), \;\; M_j^1 = \sup_{[x_{j-1},x_j]} x_{n_1}(t) \\ m_j &= \inf_{[x_{j-1},x_j]} x(t), \;\; m_j^1 = \inf_{[x_{j-1},x_j]} x_{n_1}(t) \end{aligned}$$

Using the Infimum and Supremum Tolerance Lemma, there are points s_j and t_j in $[x_{j-1}, x_j]$ so that

$$M_j - \frac{\epsilon}{5(b-a)} \; < \; x(s_j) \leq M_j \qquad (\boldsymbol{\gamma})$$

and

$$m_j \; \leq \; x(t_j) < m_j + \frac{\epsilon}{5(b-a)} \qquad (\boldsymbol{\xi})$$

Thus,

$$U(x,\boldsymbol{\pi}) - L(x,\boldsymbol{\pi}) = \sum_{\boldsymbol{\pi}} (M_j - m_j)\Delta x_j$$

The term on the right hand side can be rewritten using the standard add and subtract trick *as*

$$\sum_{\boldsymbol{\pi}} \left(M_j - x(s_j) + x(s_j) - x_{n_1}(s_j) + x_{n_1}(s_j) - x_{n_1}(t_j) + x_{n_1}(t_j) - x(t_j) + x(t_j) - m_j \right) \Delta x_j$$

We can then overestimate this term using the triangle inequality to find

$$\begin{aligned} U(x,\boldsymbol{\pi}) - L(x,\boldsymbol{\pi}) \;\leq\; & \sum_{\boldsymbol{\pi}}(M_j - x(s_j))\Delta x_j + \sum_{\boldsymbol{\pi}}(x(s_j) - x_{n_1}(s_j))\Delta x_j + \sum_{\boldsymbol{\pi}}(x_{n_1}(s_j) - x_{n_1}(t_j))\Delta x_j \\ & + \sum_{\boldsymbol{\pi}}(x_{n_1}(t_j) - x(t_j))\Delta x_j + \sum_{\boldsymbol{\pi}}(x(t_j) - m_j)\Delta x_j \end{aligned}$$

The first term can be estimated by Equation $\boldsymbol{\gamma}$ and the fifth term by Equation $\boldsymbol{\xi}$ to give

$$\begin{aligned} U(x,\boldsymbol{\pi}) - L(x,\boldsymbol{\pi}) \;<\; & \frac{\epsilon}{5(b-a)} \sum_{\boldsymbol{\pi}} \Delta x_j + \sum_{\boldsymbol{\pi}}(x(s_j) - x_{n_1}(s_j))\Delta x_j + \sum_{\boldsymbol{\pi}}(x_{n_1}(s_j) - x_{n_1}(t_j))\Delta x_j \\ & + \sum_{\boldsymbol{\pi}}(x_{n_1}(t_j) - x(t_j))\Delta x_j + \frac{\epsilon}{5(b-a)} \sum_{\boldsymbol{\pi}} \Delta x_j \end{aligned}$$

Thus,

$$\begin{aligned} U(x,\boldsymbol{\pi}) - L(x,\boldsymbol{\pi}) \;<\; & 2\frac{\epsilon}{5} + \sum_{\boldsymbol{\pi}}(x(s_j) - x_{n_1}(s_j))\Delta x_j \\ & + \sum_{\boldsymbol{\pi}}(x_{n_1}(s_j) - x_{n_1}(t_j))\Delta x_j + \sum_{\boldsymbol{\pi}}(x_{n_1}(t_j) - x(t_j))\Delta x_j \end{aligned}$$

Now apply the estimate from Equation $\boldsymbol{\alpha}$ *to the first and third terms of the equation above to conclude*

$$U(x,\boldsymbol{\pi}) \;-\; L(x,\boldsymbol{\pi}) \;\;<\;\; 4\frac{\epsilon}{5} \;+\; \sum_{\boldsymbol{\pi}} (x_{n_1}(s_j) - x_{n_1}(t_j))\Delta x_j$$

Finally, note

$$\mid x_{n_1}(s_j) - x_{n_1}(t_j) \mid \;\;\leq\;\; M_j^1 \;-\; m_j^1$$

and so

$$\begin{aligned}\sum_{\boldsymbol{\pi}} (x_{n_1}(s_j) - x_{n_1}(t_j))\Delta x_j \;\;&\leq\;\; \sum_{\boldsymbol{\pi}} (M_j^1 \;-\; m_j^1)\Delta x_j\\ &<\;\; \epsilon/5\end{aligned}$$

by Equation $\boldsymbol{\beta}$. *Thus,* $U(x,\boldsymbol{\pi}) \;-\; L(x,\boldsymbol{\pi}) < \epsilon$. *Since the partition* $\boldsymbol{\pi}$ *refining* $\boldsymbol{\pi_0}$ *was arbitrary, we see* x *satisfies the Riemann Criterion and hence, is Riemann integrable on* $[a,b]$.

It remains to show the limit interchange portion of the theorem. Since x_n *converges uniformly to* x*, given a positive* ϵ*, there is an integer* N *so that*

$$\sup_{a\leq t\leq b} \mid x_n(t) - x(t) \mid \;\;<\;\; \epsilon/(b-a), \; if\, n > N. \tag{$\boldsymbol{\zeta}$}$$

Now for any $n > N$*, we have*

$$\begin{aligned}\mid \int_a^b x(t)dt \;-\; \int_a^b x_n(t)dt \mid \;\;&=\;\; \mid \int_a^b \Big(x(t) - x_n(t)\Big)dt \mid\\ &\leq\;\; \int_a^b \Big| x(t) - x_n(t) \Big| dt\\ &\leq\;\; \int_a^b \sup_{a\leq t\leq b} \mid x_n(t) - x(t) \mid dt\\ &<\;\; \int_a^b \epsilon/(b-a)dt\\ &=\;\; \epsilon\end{aligned}$$

using Equation $\boldsymbol{\zeta}$. *This says* $\lim \int_a^b x_n(t)dt = \int_a^b x(t)dt$. ■

The next result is indispensable in modern analysis. Fundamentally, it states that a continuous real-valued function defined on a compact set can be uniformly approximated by a smooth function. This is used throughout analysis to prove results about various functions. We can often verify a property of a continuous function, f, by proving an analogous property of a smooth function that is uniformly close to f. We will only prove the result for a closed finite interval in $\Re$. The general result for a compact subset of a more general set called a *Topological Space* is a modification of this proof which is actually not that more difficult, but that is another story. We follow the development of (Simmons (5) 1963) for this proof.

Theorem 5.1.2. Weierstrass Approximation Theorem

Let f be a continuous real-valued function defined on $[0,1]$. For any $\epsilon > 0$, there is a polynomial, p, such that $|f(t) - p(t)| < \epsilon$ for all $t \in [0,1]$, that is $\| p - f \|_\infty < \epsilon$

Proof. *We first derive some equalities. We will denote the interval $[0,1]$ by I. By the binomial theorem, for any $x \in I$, we have*

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \quad = \quad (x + 1 - x)^n = 1. \qquad (\boldsymbol{\alpha})$$

Differentiating both sides of Equation $\boldsymbol{\alpha}$, we get

$$\begin{aligned} 0 &= \sum_{k=0}^{n} \binom{n}{k} \Big(k x^{k-1}(1-x)^{n-k} \; - \; x^k (n-k)(1-x)^{n-k-1} \Big) \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1} \Big(k(1-x) \; - \; x(n-k) \Big) \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1} \Big(k - nx \Big) \end{aligned}$$

Now, multiply through by $x(1-x)$, to find

$$0 \quad = \quad \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} (k - nx).$$

Differentiating again, we obtain

$$0 \quad = \quad \sum_{k=0}^{n} \binom{n}{k} \frac{d}{dx} \Big(x^k (1-x)^{n-k} (k - nx) \Big).$$

This leads to a series of simplifications. It is pretty messy and many texts do not show the details, but we think it is instructive.

$$\begin{aligned} 0 &= \sum_{k=0}^{n} \binom{n}{k} \Big[-nx^k(1-x)^{n-k} + (k-nx)\Big((k-n)x^k(1-x)^{n-k-1} + kx^{k-1}(1-x)^{n-k} \Big) \Big] \\ &= \sum_{k=0}^{n} \binom{n}{k} \Big[-nx^k(1-x)^{n-k} + (k-nx)(1-x)^{n-k-1}x^{k-1}\Big((k-n)x + k(1-x) \Big) \Big] \\ &= \sum_{k=0}^{n} \binom{n}{k} \Big(-nx^k(1-x)^{n-k} + (k-nx)^2(1-x)^{n-k-1}x^{k-1} \Big) \\ &= -n\sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k} + \sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^{k-1}(1-x)^{n-k-1} \end{aligned}$$

Thus, since the first sum is 1, we have

$$n = \sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^{k-1} (1-x)^{n-k-1}$$

and multiplying through by $x(1-x)$, we have

$$\begin{aligned} nx(1-x) &= \sum_{k=0}^{n} \binom{n}{k} (k-nx)^2 x^{k} (1-x)^{n-k} \\ \frac{x(1-x)}{n} &= \sum_{k=0}^{n} \binom{n}{k} \Big(\frac{k-nx}{n}\Big)^2 x^{k} (1-x)^{n-k} \end{aligned}$$

This last equality then leads to the

$$\sum_{k=0}^{n} \binom{n}{k} \Big(x - \frac{k}{n}\Big)^2 x^{k} (1-x)^{n-k} = \frac{x(1-x)}{n} \qquad (\boldsymbol{\beta})$$

We now define the n^{th} order Bernstein Polynomial *associated with f by*

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^{k} (1-x)^{n-k} f\Big(\frac{k}{n}\Big).$$

Note that

$$f(x) - B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^{k} (1-x)^{n-k} \Big[f(x) - f\Big(\frac{k}{n}\Big)\Big].$$

Also note that $f(0) - B_n(0) = f(1) - B_n(1) = 0$, so f and B_n match at the endpoints. It follows that

$$\mid f(x) - B_n(x) \mid \ \leq \ \sum_{k=0}^{n} \binom{n}{k} x^{k} (1-x)^{n-k} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big|. \qquad (\boldsymbol{\gamma})$$

Now, f is uniformly continuous on I since it is continuous. So, given $\epsilon > 0$, there is a $\delta > 0$ such that $|x - \frac{k}{n}| < \delta \Rightarrow |f(x) - f(\frac{k}{n})| < \frac{\epsilon}{2}$. Consider x to be fixed in $[0,1]$. The sum in Equation γ has only $n+1$ terms, so we can split this sum up as follows. Let $\{K_1, K_2\}$ be a partition of the index set $\{0, 1, ..., n\}$ such that $k \in K_1 \Rightarrow |x - \frac{k}{n}| < \delta$ and $k \in K_2 \Rightarrow |x - \frac{k}{n}| \geq \delta$. Then

$$\mid f(x) - B_n(x) \mid \leq \sum_{k \in K_1} \binom{n}{k} x^{k} (1-x)^{n-k} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big| + \sum_{k \in K_2} \binom{n}{k} x^{k} (1-x)^{n-k} \Big| f(x) - f\Big(\frac{k}{n}\Big)\Big|.$$

which implies

$$\begin{aligned}
|f(x) - B_n(x)| &\le \frac{\epsilon}{2} \sum_{k\in K_1} \binom{n}{k} x^k(1-x)^{n-k} + \sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}\Big|f(x) - f\Big(\frac{k}{n}\Big)\Big| \\
&= \frac{\epsilon}{2} + \sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}\Big|f(x) - f\Big(\frac{k}{n}\Big)\Big|.
\end{aligned}$$

Now, f is bounded on I, so there is a real number $M > 0$ such that $|f(x)| \le M$ for all $x \in I$. Hence

$$\sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}\Big|f(x) - f\Big(\frac{k}{n}\Big)\Big| \le 2M \sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}.$$

Since $k \in K_2 \Rightarrow |x - \frac{k}{n}| \ge \delta$, using Equation β, we have

$$\delta^2 \sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k} \le \sum_{k\in K_2} \binom{n}{k} \Big(x - \frac{k}{n}\Big)^2 x^k(1-x)^{n-k} \le \frac{x(1-x)}{n}.$$

This implies that

$$\sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k} \le \frac{x(1-x)}{\delta^2 n}.$$

and so combining inequalities

$$2M \sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k} \le \frac{2Mx(1-x)}{\delta^2 n}$$

We conclude then that

$$\sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}\Big|f(x) - f\Big(\frac{k}{n}\Big)\Big| \le \frac{2Mx(1-x)}{\delta^2 n}.$$

Now, the maximum value of $x(1-x)$ on I is $\frac{1}{4}$, so

$$\sum_{k\in K_2} \binom{n}{k} x^k(1-x)^{n-k}\Big|f(x) - f\Big(\frac{k}{n}\Big)\Big| \le \frac{M}{2\delta^2 n}.$$

Finally, choose n so that $n > \frac{M}{\delta^2 \epsilon}$. Then $\frac{M}{n\delta^2} < \epsilon$ implies $\frac{M}{2n\delta^2} < \frac{\epsilon}{2}$. So, Equation γ becomes

$$|\ f(x) - B_n(x)\ | \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Note that the polynomial B_n does not depend on $x \in I$, since n only depends on M, δ, and ϵ, all of which, in turn, are independent of $x \in I$. So, B_n is the desired polynomial, as it is uniformly within ϵ of f. ■

Comment 5.1.1. *A change of variable translates this result to any closed interval $[a, b]$.*

5.2 Showing Functions Are Riemann Integrable

We already know that continuous functions, monotone functions and functions of bounded variation are classes of functions which are Riemann Integrable on the interval $[a, b]$. A good reference for some of the material in this section is (Douglas (2) 1996) although it is mostly in problems and not in the text! Hence, since $f(x) = \sqrt{x}$ is continuous on $[0, M]$ for any positive M, we know f is Riemann integrable on this interval. What about the composition $\sqrt{g}$ where g is just known to be non negative and Riemann integrable on $[a, b]$? If g were continuous, since compositions of continuous functions are also continuous, we would have immediately that $\sqrt{g}$ is Riemann Integrable. However, it is not so easy to handle this case. Let's try this approach. Using Theorem 5.1.2, we know given a finite interval $[c, d]$, there is a sequence of polynomials $\{p_n(x)\}$ which converge uniformly to $\sqrt{x}$ on $[c, d]$. Of course, the polynomials in this sequence will change if we change the interval $[c, d]$, but you get the idea. To apply this here, note that since g is Riemann Integrable on $[a, b]$, g must be bounded. Since we assume g is non negative, we know that there is a positive number M so that $g(x)$ is in $[0, M]$ for all x in $[a, b]$. Thus, there is a sequence of polynomials $\{p_n\}$ which converge uniformly to $\sqrt{\cdot}$ on $[0, M]$.

Next, using Theorem 4.3.1, we know a polynomial in g is also Riemann integrable on $[a, b]$ ($f^2 = f \cdot f$ so it is integrable and so on). Hence, $p_n(f)$ is Riemann integrable on $[a, b]$. Then given $\epsilon > 0$, we know there is a positive N so that

$$| \, p_n(u) - \sqrt{u} \, | \quad < \quad \epsilon, \text{ if } n > N \text{ and } u \in [0, M].$$

Thus, in particular, since $g(x) \in [0, M]$, we have

$$| \, p_n(g(x)) - \sqrt{g(x)} \, | \quad < \quad \epsilon, \text{ if } n > N \text{ and } u \in [0, M].$$

We have therefore proved that $p_n \circ g$ converges uniformly to $\sqrt{g}$ on $[0, M]$. Then by Theorem 5.1.1, we see $\sqrt{g}$ is Riemann integrable on $[0, M]$.

If you think about it a bit, you should be able to see that this type of argument would work for any f which is continuous and g that is Riemann integrable. We state this as Theorem 5.2.1.

Theorem 5.2.1. **f Continuous and g Riemann Integrable Implies $f \circ g$ is Riemann Integrable**
If f is continuous on $g([a, b])$ where g is Riemann Integrable on $[a, b]$, then $f \circ g$ is Riemann Integrable on $[a, b]$.

Proof.

Exercise 5.2.1. *This proof is for you.*

■

In general, the composition of Riemann Integrable functions is not Riemann integrable. Here is the standard counterexample. This great example comes from (Douglas (2) 1996) . Define f on $[0, 1]$ by

$$f(y) \quad = \quad \begin{cases} 1 & \text{if } y = 0 \\ 0 & \text{if } 0 < y \leq 1 \end{cases}$$

and g on $[0,1]$ by

$$g(x) = \begin{cases} 1 & \text{if } x = 0 \\ 1/p & \text{if } x = p/q, (p,q) = 1, x \in (0,1] \text{ and } x \text{ is rational} \\ 0 & \text{if } x \in (0,1] \text{ and } x \text{ is irrational} \end{cases}$$

We see immediately that f is integrable on $[0,1]$ by Theorem 4.8.4. We can show that g is also Riemann integrable on $[0,1]$, but we will leave this as an exercise.

Exercise 5.2.2.

1. *Show g is continuous at each irrational points in $[01,]$ and discontinuous at all rational points in $[0,1]$.*

2. *Show g is Riemann integrable on $[0,1]$ with value $\int_0^1 g(x)dx = 0$.*

Now $f \circ g$ becomes

$$f(g(x)) = \begin{cases} f(1) & \text{if } x = 0 \\ f(1/p) & \text{if } x = p/q, (p,q) = 1, x \in (0,1] \text{ and } x \text{ rational} \\ f(0) & \text{if } 0 < x \leq 1 \text{ and } x \text{ irrational} \end{cases} = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if if } x \text{ rational} \in (0,1] \\ 1 & \text{if if } x \text{ irrational} \in (0,1] \end{cases}$$

The function $f \circ g$ above is not Riemann integrable as $U(f \circ g) = 1$ and $L(f \circ g) = 0$. Thus, we have found two Riemann integrable functions whose composition is not Riemann integrable!

5.3 Sets Of Content Zero

We already know the length of the finite interval $[a,b]$ is $b-a$ and we exploit this to develop the Riemann integral when we compute lower, upper and Riemann sums for a given partition. We also know that the set of discontinuities of a monotone function is countable. We have seen that continuous functions with a finite number of discontinuities are integrable and in the last section, we saw a function which was discontinuous on a countably infinite set and still was integrable! Hence, we know that a function is integrable should imply something about its discontinuity set. However, the concept of length doesn't seem to apply as there are no intervals in these discontinuity sets. With that in mind, let's introduce a new notion: the *content* of a set. We will follow the development of a set of content zero as it is done in (Sagan (4) 1974) .

Definition 5.3.1. Sets Of Content Zero

A subset S of $\Re$ is said to have content zero if and only if given any positive ϵ we can find a sequence of bounded open intervals $\{J_n^\epsilon = (a_n, b_n)\}$ either finite in number or infinite so that

$$S \subseteq \cup J_n,$$

with the total length

$$\sum (b_n - a_n) < \epsilon$$

If the sequence only has a finite number of intervals, the union and sum are written from 1 to N where N is the number of intervals and if there are infinitely many intervals, the sum and union are written from 1 to ∞.

Comment 5.3.1.

1. *A single point c in $\Re$ has content zero because $c \in (c - \epsilon/2, c + \epsilon/2)$ for all positive ϵ.*
2. *A finite number of points $S = \{c_1, \ldots, c_k\}$ in $\Re$ has content zero because $B_i = c_i \in (c_i - \epsilon/(2k), c_i + \epsilon/(2k))$ for all positive ϵ. Thus, $S \subseteq \cup_{i=1}^k B_i$ and the total length of these intervals is smaller than ϵ.*
3. *The rational numbers have content zero also. Let $\{c_i\}$ be any enumeration of the rationals. Let $B_i = (c_i - \epsilon/(2^i), c_i + \epsilon/(2^i))$ for any positive ϵ. The Q is contained in the union of these intervals and the length is smaller than $\epsilon \sum_{i=1}^\infty 1/2^i = \epsilon$.*
4. *Finite unions of sets of content zero also have content zero.*
5. *Subsets of sets of content zero also have content zero.*

Hence, the function g above is continuous on $[0, 1]$ except on a set of content zero. We make this more formal with a definition.

Definition 5.3.2. Continuous Almost Everywhere

The function f defined on the interval $[a, b]$ is said to be continuous almost everywhere if the set of discontinuities of f has content zero. We abbreviate the phrase almost everywhere *by writing a.e.*

We are now ready to prove an important theorem which is known as the **Riemann - Lebesgue Lemma**. This is also called **Lebesgue's Criterion For the Riemann Integrability of Bounded Functions** . We follow the proof given in (Sagan (4) 1974) .

Theorem 5.3.1. Riemann - Lebesgue Lemma

(i) $f \in B[a, b]$ and continuous a.e. implies $f \in RI[a, b]$.

(ii) $f \in RI[a, b]$ implies f is continuous a.e.

Proof. *The proof of this result is fairly complicated. So grab a cup of coffee, a pencil and prepare for a long battle!*
(i):

Subproof. *We will prove this by showing that for any positive ϵ, we can find a partition $\boldsymbol{\pi_0}$ so that the Riemann Criterion is satisfied. First, since f is bounded, there is are numbers m and M so that $m \leq f(x) \leq M$ for all x in $[a, b]$. If m and M we the same, then f would be constant and it would therefore be continuous. If this case, we know f is integrable. So we can assume without loss of generality that $M - m > 0$. Let D denote the set of points in $[a, b]$ where f is not continuous. By assumption, the content of D is zero. Hence, given a positive ϵ there is a sequence of bounded open intervals $J_n = (a_n, b_n)$ (we will assume without loss of generality that there are infinitely many such intervals) so that*

$$D \subseteq \cup J_n, \ \sum (b_n - a_n) < \epsilon/(2(M-m)).$$

Now if x is from $[a, b]$, x is either in D or in the complement of D, D^C. Of course, if $x \in D^C$, then f is continuous at x. The set

$$E = [a, b] \cap \left(\cup J_n \right)^C$$

is compact and so f must be uniformly continuous on E. Hence, for the ϵ chosen, there is a $\delta > 0$ so that

$$\mid f(y) - f(x) \mid < \epsilon/(8(b-a)), \qquad (*)$$

if $y \in (x - \delta, x + \delta) \cap E$. Next, note that

$$\boldsymbol{O} = \{J_n, B_{\delta/2}(x) \mid x \in E\}$$

is an open cover of $[a, b]$ and hence must have a finite sub cover. Call this finite sub cover $\boldsymbol{O'}$ and label its members as follows:

$$\boldsymbol{O'} = \{J_{n_1}, \ldots, J_{n_r}, B_{\delta/2}(x_1), \ldots, B_{\delta/2}(x_s)\}$$

Then it is also true that we know that

$$[a, b] \subseteq \boldsymbol{O''} \ = \ \{J_{n_1}, \ldots, J_{n_r}, B_{\delta/2}(x_1) \cap E, \ldots, B_{\delta/2}(x_s) \cap E\}$$

All of the intervals in $\boldsymbol{O''}$ have endpoints. Throw out any duplicates and arrange these endpoints in increasing order in $[a, b]$ and label them as $y_1, \ldots, y_{p-1}$. Then, let

$$\boldsymbol{\pi_0} = \{y_0 = a, y_1, y_2, \ldots, y_{p-1}, y_p = b\}$$

be the partition formed by these points. Recall where the points y_j come from. The endpoints of the $B_{\delta/2}(x_i) \cap E$ sets are not in any of the intervals J_{n_k}. So suppose two successive points y_{j-1} and y_j satisfied y_{j-1} is in an interval J_{n_k} and the next point y_j was an endpoint of a $B_{\delta/2}(x_i) \cap E$ set which is also inside J_{n_k}. By our construction, this can not happen as all of the $B_{\delta/2}(x_i) \cap E$ are disjoint from the J_{n_k} sets. Hence, the next point y_j either must be in the set J_{n_k} also or it must be outside. If y_{j-1} is inside and y_j is outside, this is also a contradiction as this would give us a third point, call it z temporarily, so that

$$y_{j-1} \ < \ z \ < y_j$$

with z a new distinct endpoint of the finite cover $\boldsymbol{O''}$*. Since we have already ordered these points, this third point is not a possibility. Thus, we see* (y_{j-1}, y_j) *is in some* J_{n_k} *or neither of the points is in any* J_{n_k}*. Hence, we have shown that given the way the points* y_j *were chosen, either* (y_{j-1}, y_j) *is inside some interval* J_{n_q} *or it's closure* $[y_{j-1}, y_j]$ *lies in none of the* J_{n_q} *for any* $1 \leq q \leq r$*. But that means* (y_{j-1}, y_j) *lies in some* $\hat{B}_{\delta/2}(x_i)$*. Note this set uses the radius* $\delta/2$ *and so we can say the closed interval* $[y_{j-1}, y_j]$ *must be contained in some* $\hat{B}_\delta(x_i)$*.*

Now we separate the index set $\{1, 2, \ldots, p\}$ *into two disjoint sets. We define* A_1 *to be the set of all indices* j *so that* (y_{j-1}, y_j) *is contained in some* J_{n_k}*. Then we set* A_2 *to be the complement of* A_1 *in the entire index set, i.e.* $A_2 = \{1, 2, \ldots, p \ - A_1$*. Note, by our earlier remarks, if* j *is in* A_2*,* $[y_{j-1}, y_j]$ *is contained in some* $B_\delta(x_i) \cap E$*. Thus,*

$$\begin{aligned} U(f, \boldsymbol{\pi_0}) - L(f, \boldsymbol{\pi_0}) &= \sum_{j=1}^{n} \Bigg(M_j - m_j \Bigg) \Delta y_j \\ &= \sum_{j \in A_1} \Bigg(M_j - m_j \Bigg) \Delta y_j \ + \sum_{j \in A_2} \Bigg(M_j - m_j \Bigg) \Delta y_j \end{aligned}$$

Let's work with the first sum: we have

$$\begin{aligned} \sum_{j \in A_1} \Bigg(M_j - m_j \Bigg) \Delta y_j &\leq \Bigg(M - m \Bigg) \sum_{j \in A_1} \Delta y_j \\ &< (M - m)\, \epsilon / (2(M - m)) \ = \ \epsilon/2 \end{aligned}$$

Now if j *is in* A_2*, then* $[y_{j-1}, y_j]$ *is contained in some* $B_\delta(x_i) \cap E$*. So any two points* u *and* v *in* $[y_{j-1}, y_j]$ *satisfy* $\mid u - x_i \mid < \delta$ *and* $\mid v - x_i \mid < \delta$*. Since these points are this close, the uniform continuity condition, Equation* $*$*, holds. Therefore*

$$\mid f(u) - f(v) \mid \quad \leq \quad \mid f(u) - f(x_i) \mid + \mid f(v) - f(x_i) \mid < \ \epsilon/(4(b - a)).$$

This holds for any u *and* v *in* $[y_{j-1}, y_j]$*. In particular, we can use the Supremum and Infimum Tolerance Lemma to choose* u_j *and* v_j *so that*

$$M_j - \epsilon/(8(b - a)) < f(u_j), \ \ m_j + \epsilon/(8(b - a)) > f(v_j).$$

It then follows that

$$M_j - m_j < f(u_j) - f(v_j) + \epsilon/(4(b - a)).$$

Now, we can finally estimate the second summation term. We have

$$\begin{aligned}
\sum_{j \in A_2} \Bigg(M_j - m_j \Bigg) \Delta y_j &< \sum_{j \in A_2} \Bigg(\mid f(u_j) - f(v_j) \mid + \epsilon/(4(b-a)) \Bigg) \Delta y_j \\
&< \sum_{j \in A_2} \Bigg(\mid f(u_j) - f(v_j) \mid \Bigg) \Delta y_j \; + \; \epsilon/(4(b-a)) \sum_{j \in A_2} \Delta y_j \\
&< \epsilon/(4(b-a)) \sum_{j \in A_2} \Delta y_j \; + \; \epsilon/(4(b-a)) \sum_{j \in A_2} \Delta y_j \\
&< \epsilon/2
\end{aligned}$$

Combining our estimates, we have

$$\begin{aligned}
U(f, \boldsymbol{\pi_0}) - L(f, \boldsymbol{\pi_0}) &= \sum_{j \in A_1} \Bigg(M_j - m_j \Bigg) \Delta y_j \; + \; \sum_{j \in A_2} \Bigg(M_j - m_j \Bigg) \Delta y_j \\
&< \epsilon/2 + \epsilon/2 \; = \; \epsilon.
\end{aligned}$$

Any partition $\boldsymbol{\pi}$ that refines $\boldsymbol{\pi_0}$ will also satisfy $U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) < \epsilon$. Hence, f satisfies the Riemann Criterion and so f is integrable. □

(ii):

Subproof. *We begin by noting that if f is discontinuous at a point x in $[a, b]$, if and only if there is a positive integer m so that*

$$\forall \delta > 0, \; \exists y \in (x - \delta, x + \delta) \cap [a, b] \quad \ni \quad \mid f(y) - f(x) \mid \geq 1/m.$$

This allows us to define some interesting sets. Define the set E_m by

$$E_m \quad = \quad \{ x \in [a, b] \mid \forall \delta > 0 \; \exists y \in (x - \delta, x + \delta) \cap [a, b] \; \ni \; \mid f(y) - f(x) \mid \geq 1/m, \}$$

Then, the set of discontinuities of f, $\boldsymbol{D}$ can be expressed as $\boldsymbol{D} = \cup_{j=1}^{\infty} E_m$.

Now let $\boldsymbol{\pi} = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a, b]$. Then, given any positive integer m, the open subinterval $[x_{k-1}, x_k]$ either intersects E_m or it does not. Define

$$\begin{aligned}
A_1 &= \Bigg\{ k \in \{1, \ldots, n\} \mid (x_{k-1}, x_k) \cap E_m \neq \emptyset \Bigg\}, \\
A_2 &= \Bigg\{ k \in \{1, \ldots, n\} \mid (x_{k-1}, x_k) \cap E_m = \emptyset \Bigg\}
\end{aligned}$$

By construction, we have $A_1 \cap A_2 = \emptyset$ and $A_1 \cup A_2 = \{1, \ldots, n\}$.
We assume f is integrable on $[a, b]$. So, by the Riemann Criterion, given $\epsilon > 0$, and a positive integer m, there is a partition $\boldsymbol{\pi_0}$ such that

$$U(f, \boldsymbol{\pi}) - L(f, \boldsymbol{\pi}) < \epsilon/(2m), \; \forall \boldsymbol{\pi_0} \preceq \boldsymbol{\pi}. \qquad (**)$$

It follows that if $\boldsymbol{\pi_0} = \{y_0, y_1, \ldots, y_n\}$, *then*

$$
\begin{aligned}
U(f, \boldsymbol{\pi_0}) - L(f, \boldsymbol{\pi_0}) &= \sum_{k=1}^{n} (M_k - m_k)\Delta y_k \\
&= \sum_{k \in A_1} (M_k - m_k)\Delta y_k + \sum_{k \in A_2} (M_k - m_k)\Delta y_k
\end{aligned}
$$

If k *is in* A_1, *then by definition, there is a point* u_k *in* E_m *and a point* v_k *in* (y_{k-1}, y_k) *so that* $| f(u_k) - f(v_k) | \geq 1/m$. *Also, since* u_k *and* v_k *are both in* (y_{k-1}, y_k),

$$
M_k - m_k \geq | f(u_k) - f(v_k) | .
$$

Thus,

$$
\sum_{k \in A_1} (M_k - m_k)\Delta y_k \;\geq\; \sum_{k \in A_1} | f(u_k) - f(v_k) | \Delta y_k \;\geq\; (1/m) \sum_{k \in A_1} \Delta y_k.
$$

Also, the second term, $\sum_{k \in A_2} (M_k - m_k)\Delta y_k$ *is non-negative and so using Equation* $**$, *we find*

$$
\epsilon/(2m) \;>\; U(f, \boldsymbol{\pi_0} - L(f, \boldsymbol{\pi_0} \;\geq\; (1/m) \sum_{k \in A_1} \Delta y_k.
$$

which implies $\sum_{k \in A_1} \Delta y_k < \epsilon/2$.

The partition $\boldsymbol{\pi_0}$ *divides* $[a, b]$ *as follows:*

$$
\begin{aligned}
[a, b] &= \Big(\cup_{k \in A_1} (y_{k-1}, y_k) \Big) \cup \Big(\cup_{k \in A_2} (y_{k-1}, y_k) \Big) \cup \Big(\{y_0, \ldots, y_n\} \Big) \\
&= C_1 \cup C_2 \cup \boldsymbol{\pi_0}
\end{aligned}
$$

By the way we constructed the sets E_m, *we know* E_m *does not intersect* C_2. *Hence, we can say*

$$
E_m = \Big(C_1 \cap E_m \Big) \cup \Big(E_m \cap \boldsymbol{\pi_0} \Big)
$$

Therefore, we have $C_1 \cap E_m \subseteq \cup_{k \in A_1} (y_{k-1}, y_k)$ *with* $\sum_{k \in A_1} \Delta y_k < \epsilon/2$. *Since* ϵ *is arbitrary, we see* $C_1 \cap E_m$ *has content zero. The other set* $E_m \cap \boldsymbol{\pi_0}$ *consists of finitely many points and so it also has content zero by the comments at the end of Definition 5.3.1. This shows that* E_m *has content zero since it is the union of two sets of content zero. We finish by noting* $D = \cup E_m$ *also has content zero. The proof of this we leave as an exercise.*

Exercise 5.3.1. *Prove that if* $F_n \subseteq [a, b]$ *has content zero for all* n, *then* $F = \cup F_n$ *also has content zero.*

□

■

Chapter 6

Cantor Set Experiments

We now begin a series of personal investigations into the construction of an important subset of $[0, 1]$ called the Cantor Set. We follow a great series of homework exercise outlined, without solutions, in a really hard but extraordinarily useful classical analysis text by Stromberg, (Stromberg (6) 1981) .

6.1 The Generalized Cantor Set

Let (a_n) for $n \geq 0$ be a fixed sequence of real numbers which satisfy

$$a_0 \quad = \quad 1,\ 0 < 2a_n < a_{n-1} \tag{6.1}$$

Define the sequence (d_n) by

$$d_n = a_{n-1} - 2a_n$$

Note each $d_n > 0$. We can use the sequence (a_n) to define a collection of intervals $J_{n,k}$ and $I_{n,k}$ as follows.

(0) $J_{0,1} = [0, 1]$ which has length a_0.

(1) $J_{1,1} = [0, a_1]$ and $J_{1,2} = [1 - a_1, 1]$. You can see each of these intervals has length a_1. We let $W_{1,1} = J_{1,1} \cup J_{1,2}$ and $I_{1,1} = J_{0,1} - W_{1,1}$ where the *minus* symbol used here represents set difference. This step creates an open interval of $[0, 1]$ which has length $d_1 > 0$. Let $P_1 = J_{1,1} \cup J_{1,2}$. This is a closed set.

(2) Set $J_{2,1} = [0, a_2]$, $J_{2,2} = [a_1 - a_2, a_1]$, $J_{2,3} = [1 - a_1, 1 + a_2 - a_1]$, and $J_{2,4} = [1 - a_2, 1]$. These 4 closed subintervals have length a_2. It is not so mysterious how we set up the $J_{2,k}$ intervals. Step (1) created a closed interval $[0, a_1]$, an open interval $(a_1, 1 - a_1)$ and another closed interval $[1 - a_1, 1]$. The first closed subinterval is what we have called $J_{1,1}$. Divide it into three parts; the

first part will be a closed interval that starts at the beginning of $J_{1,1}$ and has length a_2 and the third part will be closed interval of length a_2 that ends at the last point of $J_{1,1}$. When these two closed intervals are subtracted from $J_{1,1}$, an open interval will remain. The length of $J_{1,1}$ is a_1. So the open interval must have length $a_1 - 2a_2 = d_2$. A little thought tells us that the first interval must be $[0, a_2]$ (which we have named $J_{2,1}$) and the third interval must be $[a_1 - a_2, a_1]$ (which we have named $J_{2,2}$). To get the intervals $J_{2,3}$ and $J_{2,4}$, we divide $J_{1,2}$ into the same type of three subintervals as we did for $J_{1,1}$. The first and third must have length a_2 which will give an open interval in the inside of length d_2. This will give $J_{2,3} = [1 - a_1, 1 - a_1 + a_2]$ and $j_{2,4} = [1 - a_2, 1]$. Then let $W_{2,1} = J_{2,1} \cup J_{2,2}$, and $W_{2,2} = J_{2,3} \cup J_{2,4}$. Then create new intervals by letting $I2, 1 = J_{1,1} - W_{2,1}$ and $I2, 2 = J_{1,2} - W_{2,2}$. We have now created 4 open subintervals of length d_2. Let $P_2 = J_{2,1} \cup J_{2,2} \cup J_{2,3} \cup J_{2,4}$. We can write this more succinctly as $P_2 = \cup\{J_{2,k} | 1 <= k <= 2^2\}$. Again, notice that P_2 is a closed set that consists of 4 closed subintervals of length a_2.

Let's look even more closely at the details. A careful examination of the process above with pen and paper in hand gives the following table that characterizes the left hand endpoint of each of the intervals $J_{2,k}$.

$J_{2,1}$	0
$J_{2,2}$	$a_2 + d_2$
$J_{2,3}$	$2a_2 + d_2 + d_1$
$J_{2,4}$	$3a_2 + 2d_2 + d_1$

Since we know the left hand endpoint and the length is always a_2, this fully characterizes the subintervals $J_{2,k}$. Also, as a check, the last endpoint $3a_2 + 2d_2 + d_1$ plus one more a_2 should add up to 1. We find

$$\begin{aligned} 4a_2 + 2d_2 + d_1 &= 4a_2 + 2(a_1 - 2a_2) + (a_0 - 2a_1) \\ &= a_0 \;=\; 1. \end{aligned}$$

(3) Step (2) has created 4 closed subintervals $J_{2,k}$ of length a_2 and 2 new open intervals $I_{2,i}$ of length d_2. There is also the first open interval $I_{1,1}$ of length d_1 which was abstracted from $[0, 1]$. Now we repeat the process described in Step (2) on each closed subinterval $J_{2,k}$. We do not need to use the auxiliary sets $W_{3,i}$ now as we can go straight into the subdivision algorithm. We divide each of these intervals into 3 pieces. The first and third will be of length a_3. This leaves an open interval of length d_3 between them. We label the new closed subintervals so created by $J_{3,k}$ where k now ranges from 1 to 8. The new intervals have left hand endpoints

$J_{3,1}$	0
$J_{3,2}$	$a_3 + d_3$
$J_{3,3}$	$2a_3 + d_3 + d_2$
$J_{3,4}$	$3a_3 + 2d_3 + d_2$
$J_{3,5}$	$4a_3 + 2d_3 + d_2 + d_1$
$J_{3,6}$	$5a_3 + 3d_3 + d_2 + d_1$
$J_{3,7}$	$6a_3 + 3d_3 + 2d_2 + d_1$
$J_{3,8}$	$7a_3 + 4d_3 + 2d_2 + d_1$

Each of these subintervals have length a_3 and a simple calculation shows $(7a_3+4d_3+2d_2+d_1)+a_3 = 1$ as desired. There are now 4 more open intervals $I_{3,i}$ giving a total of 6 open subintervals arranged as follows:

	Parent	Length
$I_{1,1}$	$J_{0,1}$	d_1
$I_{2,1}$	$J_{1,1}$	d_2
$I_{2,2}$	$J_{1,2}$	d_2
$I_{3,1}$	$J_{2,1}$	d_3
$I_{3,2}$	$J_{2,2}$	d_3
$I_{3,3}$	$J_{2,3}$	d_3
$I_{3,4}$	$J_{2,4}$	d_4

We define $P_3 = \cup\{J_{3,k} | 1 <= k <= 2^3\}$ and note that $P_1 \cap P_2 \cap P_3 = P_3$.

We can, of course, continue this process recursively. Thus, after Step n, we have constructed 2^n closed subintervals $J_{n,k}$ each of length a_n. The union of these subintervals is labeled P_n and is therefore defined by $P_n = \cup\{J_{n,k} | 1 <= k <= 2^n\}$. The left hand endpoints of $J_{n,k}$ can be written in a compact and illuminating form, but we will delay working that out until later. Now, we can easily see the form of the left hand endpoints for the first few intervals:

$J_{n,1}$	0
$J_{n,2}$	$a_n + d_n$
$J_{n,3}$	$2a_n + d_n + d_{n-1}$
$J_{n,4}$	$3a_n + 2d_n + d_{n-1}$

Definition 6.1.1. The Generalized Cantor Set
Let (a_n), $N \geq 0$ satisfy Equation 6.1. We call such a sequence a Cantor Set Generating Sequence and we define the Cantor Set generated by (a_n) to be the set $P = \cap_{n=1}^{\infty} P_n$, where the sets P_n are defined recursively via the discussion in this section. We will denote the generalized Cantor Set generated by the Cantor Sequence (a_n) by $\boldsymbol{C_a}$.

Comment 6.1.1. *The Cantor Set generated by the sequence $(1/3^n)$, $n \geq 0$ is very famous and is called the* Middle Thirds *set because we are always removing the middle third of each interval in the construction process. We will denote the Middle Third Cantor set by $\boldsymbol{C}$.*

Exercise 6.1.1. *Write out the explicit endpoints of all these intervals up to and including Step 4. Illustrate this process with clearly drawn tables and graphs.*

Exercise 6.1.2. *Write out explicitly P_1, P_2, P_3 and P_4. Illustrate this process with clearly drawn tables and graphs.*

Exercise 6.1.3. *Do the above two steps for the choice $a_n = 3^{-n}$ for $n >= 0$. Illustrate this process with clearly drawn tables and graphs.*

Exercise 6.1.4. *Do the above two steps for the choice $a_n = 5^{-n}$ for $n >= 0$. Illustrate this process with clearly drawn tables and graphs.*

Exercise 6.1.5. *As mentioned, the above construction process above can clearly be handled via induction. Prove the following:*

(a) $P_{n-1} - P_n = \cup\{I_{n,k} \,|\, 1 <= k <= 2^{n-1}\}$

(b) Let $P = \cap_{n=0}^{\infty} P_n$*. Then* $P_0 - P = \cup_{n=1}^{\infty} \left(P_{n-1} - P_n\right)$

6.2 Representing The Generalized Cantor Set

We are now in a position to prove additional properties about the Cantor Set $\boldsymbol{C_a}$ for a Cantor generating sequence (a_n). Associate with (a_n) the sequence (r_n) whose entries are defined by $r_n = a_{n-1} - a_n$. Let S denote the set of all sequences of real numbers whose values are either 0 or 1; i.e. $S = \{x = (x_n) \,|\, x_n = 0 \text{ or } x_n = 1\}$. Now define the mapping $f : S \to \boldsymbol{C_a}$ by

$$f(x) \;=\; \sum_{n=1}^{\infty} x_n r_n \tag{6.2}$$

Theorem 6.2.1. Representing The Cantor Set

1. *f is well - defined.*
2. *$f(x)$ is an element of $\boldsymbol{C_a}$.*
3. *f is $1-1$ from S to $\boldsymbol{C_a}$.*
4. *f is onto $\boldsymbol{C_a}$.*

Proof. *You will prove this Theorem by establishing a series of results.*

Exercise 6.2.1. *For any Cantor generating sequence (a_n), we have $\lim_n a_n = 0$.*

Exercise 6.2.2. *Show*

$$\sum_{j=n+1}^{\infty} r_j \;=\; \lim_m \sum_{j=n+1}^{m} r_j \;=\; \lim_m (a_n - a_m) \;=\; a_n$$

Exercise 6.2.3. $r_n > \sum_{j=n+1}^{\infty} r_j$.

Exercise 6.2.4. *For $n >= 1$ and any finite sequence $(x_1, x_2, ..., x_n)$ of 0's and 1's, define the closed interval*

$$J(x_1, ..., x_n) \;=\; [\sum_{j=1}^{n} x_j r_j, a_n + \sum_{j=1}^{n} x_j r_j]$$

Show

1. *Show* $J(0) = [0, a_1] = J_{1,1}$.
2. *Show* $J(1) = [r_1, a_1 + r_1] = [1 - a_1, 1] = J_{1,2}$.
3. *Now use induction on* n *to show that the intervals* $J(x_1, ..., x_n)$ *are exactly the* 2^n *intervals* $J_{n,k}$ *for* $1 <= k <= 2^n$ *that we described in the previous section.*

   Hint. *i.e. assume true for* $n-1$. *Then we can assume that there is a unique* $(x_1, ..., x_{n-1})$ *choice so that* $J_{n-1,k} = J(x_1, ..., x_{n-1})$.

   Recall how the J's are constructed. At Step $n-1$, *the interval* $J_{n-1,k}$ *is used to create* 2 *more intervals on level* n *by removing a piece. The* 2 *intervals left both have length* a_n *and we would denote them by* $J_{n,2k-1}$ *and* $J_{n,2k}$. *Now use the definition of the closed intervals* $J(x_1, ..., x_n)$ *to show that (remember our* $x_1, ..., x_{n-1}$ *are fixed)*

$$
\begin{aligned}
J(x_1, ..., x_{n-1}, 0) &= J_{n,2k-1} \\
J(x_1, ..., x_{n-1}, 1) &= J_{n,2k}
\end{aligned}
$$

   This will complete the induction. □

Exercise 6.2.5. *Let* x *be in* S. *Show that* $f(x)$ *is in* $J(x_1, ..., x_n)$ *for each* n.
Sketch Of Argument: We know that each $J(x_1, ..., x_n) = J_{n,k}$ *for some* k. *Let this* k *be written* $k(x, n)$ *to help us remember that it depends on the* x *and the* n. *Also remember that* $1 <= k(x, n) <= 2^n$. *So* $f(x)$ *is in* $J_{n,k(x,n)}$ *which is contained in* P_n *Hence,* $f(x)$ *is in* P_n *for all* n *which shows f(x) is in* $\boldsymbol{C_a}$. *This shows* f *maps* S *into* $\boldsymbol{C_a}$.

Exercise 6.2.6. *Now let* x *and* y *be distinct in* S. *Choose an index* j *so that* x_j *is different from* y_j. *Show this implies that* $f(x)$ *and* $f(y)$ *then belong to different closed intervals on the* j^{th} *level. This implies* $f(x)$ *is not the same as* $f(y)$ *and so* f *is* $1-1$ *on* S.

Exercise 6.2.7. *Show* f *is surjective. To do this, let* z *be in* $\boldsymbol{C_a}$. *Since* z *is in* P_1, *either* z *is in* $J(0)$ *or* z *is in* $J(1)$. *Choose* x_1 *for that* z *is in* $J(x_1)$. *Then assuming* $x_1, ..., x_{n-1}$ *have been chosen, we have* z *is in* $J(x_1, ..., x_{n-1})$. *Now* z *is in*

$$P_n \cap J(x_1, ..., x_{n-1}) = J(x_1, ..., x_{n-1}, 0) \cup J(x_1, ..., x_{n-1}, 1).$$

This tells us how to choose x_n.

Hence, by induction, we can find a sequence (x_n) *in* S *so that* z *is in intersection over* n *of* $J(x_1, ..., x_n)$. *But by our earlier arguments,* $f(x)$ *is in the same intersection!*

Finally, each of these closed intervals has length a_n *which we know goes to* 0 *in the limit on* n. *So* z *and* $f(x)$ *are both in a decreasing sequence of sets whose lengths go to* 0. *Hence* z *and* $f(x)$ *must be the same. (This uses what is called the Cantor Intersection Theorem).*

■

We can also prove a result about the internal structure of the generalized Cantor set: it can not contain any open intervals.

Exercise 6.2.8. *Prove* $\boldsymbol{C_a}$ *contains no open intervals.*

In addition, we have the following result:

Exercise 6.2.9. *The limit of* $2^n a_n$ *always exists and is in* $[0, 1]$.

6.3 The Cantor Function

We now prove additional interesting results that arise from the use of generalized Cantor sets via a series of exercises that you complete. As usual, let (a_n) be a Cantor Set generating sequence. Using the function f defined in the previous section, let's define the mapping ϕ by

$$\phi((x_n)) \quad = \quad \sum_{j=1}^{\infty} x_j \, (1/2^j)$$

Hence, $\phi : S \to [0, 1]$. and $\phi \circ f : S \to [0, 1]$. Let the mapping $\Psi = \phi \circ f^{-1}$. Note $\Psi : \boldsymbol{C_a} \to [0, 1]$.

Exercise 6.3.1. ϕ *maps* S *one to one and onto* $[0, 1]$ *with a suitable restriction on the base* 2 *representation of a number in* $[0, 1]$.

Exercise 6.3.2. $x < y$ *in* $\boldsymbol{C_a}$ *implies* $\Psi(x) \leq \Psi(y)$.

Exercise 6.3.3. $\Psi(x) = \Psi(y)$ *if and only if* (x, y) *is one of the intervals removed in the Cantor set construction process, i.e.*

$$(x, y) \quad = \quad \left(\sum_{j=1}^{n-1} x_j r_j + a_n, \sum_{j=1}^{n-1} x_j r_j + r_n \right)$$

Exercise 6.3.4. *In the case where* $\Psi(x) = \Psi(y)$ *extend the mapping* Ψ *to* $[0, 1] - \boldsymbol{C_a}$ *by*

$$\Psi(t) = \Psi(x) = \Psi(y), \; x < t < y.$$

Finally, define $\Psi(0) = 0$ *and* $\Psi(1) = 1$. *Prove* $\Psi : [0, 1] \to [0, 1]$ *is a non increasing continuous map of* $[0, 1]$ *onto* $[0, 1]$ *and is constant on each* component interval *of* $[0, 1] - \boldsymbol{C_a}$ *where* component interval *means the* $I_{n,k}$ *sets we constructed in the Cantor set construction process.*

Comment 6.3.1. *If* $\boldsymbol{C_a}$ *is the Cantor set constructed from the sequence* $(1/3^n)$, *we call* Ψ *the* Lebesgue Singular Function.

Now, let $\boldsymbol{C}$ be a Cantor set constructed from the generating sequence (a_n) where $\lim 2^n a_n = 0$. Let Ψ be the mapping discussed above for this $\boldsymbol{C}$. Define the mapping $g : [0, 1] \to [0, 1]$ by $g(x) = (\Psi(x) + x)/2$.

Exercise 6.3.5. *Prove* g *is strictly increasing and continuous from* $[0, 1]$ *onto* $[0, 1]$.

Exercise 6.3.6. *Prove that*

$$g(\sum_{j=1}^{\infty} x_j \, r_j) = \sum_{j=1}^{\infty} x_j \, r_j'$$

where $r_j\prime = (1/2^j + r_j)/2$.

Exercise 6.3.7. *Prove* $\boldsymbol{C'} = g(\boldsymbol{C})$ *is also a generalized Cantor set.*

Comment 6.3.2. *Note that the sequence* $a'_j = (1/2)(1/2^j + a_j)$ *is also a Cantor generating sequence that gives the desired* $r_j\prime$ *for the previous exercise.*

Exercise 6.3.8. *Compute the content of the Cantor set generated by* a_n *when* $\lim 2^n a_n = 0$ *and also the content of the Cantor set* $\boldsymbol{C'} = g(\boldsymbol{C})$.

In later chapters, this function g will be of great importance!

Chapter 7

The Riemann-Stieltjes Integral

In classical analysis, the Riemann-Stieltjes integral was the first attempt to generalize the idea of the size, or measure, of a subset of the real numbers. Instead of simply using the length of an interval as a measure, we can use any function that satisfies the same properties as the length function.

Let f and g be any bounded functions on the finite interval $[a, b]$. If $\boldsymbol{\pi}$ is any partition of $[a, b]$ and $\boldsymbol{\sigma}$ is any evaluation set, we can extend the notion of the Riemann sum $S(f, \boldsymbol{\pi}, \boldsymbol{\sigma}$ to the more general **Riemann - Stieljes** sum as follows:

Definition 7.0.1. The Riemann - Stieljes Sum
Let $f, g \in B[a, b]$, $\boldsymbol{\pi} \in \Pi[a, b]$ and $\boldsymbol{\sigma} \subseteq \boldsymbol{\pi}$. Let the partition points in $\boldsymbol{\pi}$ be $\{x_0, x_1, \ldots, x_p\}$ and the evaluation points be $\{s_1, s_2, \ldots, s_p\}$ as usual. Define

$$\Delta g_j = g(x_j) - g(x_j - i),\ 1 \ \leq \ j \ \leq \ p.$$

and the Riemann - Stieljes sum for **integrand** *f and* **integrator** *g for partition $\boldsymbol{\pi}$ and evaluation set $\boldsymbol{\pi}$ by*

$$S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma}) = \sum_{j \in \boldsymbol{\pi}} f(s_j)\, \Delta g_j$$

This is also called the **Riemann - Stieljes sum for the function f with respect to the function** *g for partition $\boldsymbol{\pi}$ and evaluation set $\boldsymbol{\sigma}$.*

Of course, you should compare this definition to Definition 4.1.1 to see the differences! We can then define the **Riemann - Stieljes** integral of f with respect to g using language very similar to that of Definition 4.1.2.

Definition 7.0.2. The Riemann - Stieljes Integral
Let $f, g \in B[a,b]$. If there is a real number I so that for all positive ϵ, there is a partition $\boldsymbol{\pi_0} \in \Pi[a,b]$ so that

$$\left| S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma}) - I \right| < \epsilon$$

for all partitions $\boldsymbol{\pi}$ that refine $\boldsymbol{\pi_0}$ and evaluation sets $\boldsymbol{\sigma}$ from $\boldsymbol{\pi}$, then we say f is Riemann - Stieljes integrable with respect to g on $[a,b]$. We call the value I the Riemann - Stieljes integral of f with respect to g on $[a,b]$. We use the symbol

$$I = RS(f, g; a, b)$$

to denote this value. We call f the **integrand** *and g the* **integrator**.

As usual, there is the question of what pairs of functions (f, g) will turn out to have a finite Riemann - Stieljes integral. The collection of the functions f from $B[a,b]$ that are Riemann - Stieljes integrable with respect to a given integrator g from $B[a,b]$ is denoted by $RS[g, a, b]$.

Comment 7.0.3. *If $g(x) = x$ on $[a,b]$, then $RS[g,a,b] = RI[a,b]$ and $RS(f,g;a,b) = \int_a^b f(x)dx$.*

Comment 7.0.4. *We will use the standard conventions: $RS(f,g;a,b) = -RS(f,g;b,a)$ and $RS(f,g;a;a) =$ 0.*

7.1 Standard Properties Of The Riemann - Stieljes Integral

We can easily prove the usual properties that we expect an *integration* type mapping to have.

Theorem 7.1.1. The Linearity of the Riemann - Stieljes Integral
If f_1 and f_2 are in $RS[g,a,b]$, then

(i)

$$c_1 f_1 + c_2 f_2 \in RS[g,a,b], \ \forall c_1, c_2 \in \Re$$

(ii)

$$RS(c_1 f_1 + c_2 f_2, g; a, b) = c_1 RS(f_1, g; a, b) \ + \ c_2 RS(f_2, g; a, b)$$

If $f \in RS[g_1,a,b]$ and $f \in RS[g_2,a,b]$ then

(i)

$$f \in RS[c_1 g_1 + c_2 g_2, a, b], \ \forall c_1, c_2 \in \Re$$

(ii)

$$RS(f, c_1 g_1 + c_2 g_2; a, b) = c_1 RS(f, g_1; a, b) \ + \ c_2 RS(f, g_2; a, b)$$

Proof.

Exercise 7.1.1. *We leave these proofs to you as an exercise.*

The proof of these statements is quite similar in spirit to those of Theorem 4.1.1. You should compare the techniques! ■

To give you a feel for the kind of partition arguments we use for Riemann - Stieljes proofs (you will no doubt enjoy working out these details for yourselves in various exercises), we will go through the proof of the standard **Integration By Parts** formula in this context.

Theorem 7.1.2. Riemann Stieljes Integration By Parts

If $f \in RS[g, a, b]$, then $g \in RS[f, a, b]$ and

$$RS(g, f; a, b) = f(x)g(s)\Big|_a^b - RS(f, g; a, b)$$

Proof. *Since $f \in RS[g, a, b]$, there is a number $I_f = RS(f, g; a, b)$ so that given a positive ϵ, there is a partition $\boldsymbol{\pi_0}$ such that*

$$\left| S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma} - I_f \right| \quad < \quad \epsilon, \; \boldsymbol{\pi_0} \preceq \boldsymbol{\pi}, \; \boldsymbol{\sigma} \subseteq \boldsymbol{\pi}. \qquad (\boldsymbol{\alpha})$$

For such a partition $\boldsymbol{\pi}$ and evaluation set $\boldsymbol{\sigma} \subseteq \boldsymbol{\pi}$, we have

$$\begin{aligned} \boldsymbol{\pi} &= \{x_0, x_1, \ldots, x_p\}, \\ \boldsymbol{\sigma} &= \{s_1, \ldots, s_p\} \end{aligned}$$

and

$$S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \quad = \quad \sum_{\boldsymbol{\pi}} g(s_j) \Delta f_j.$$

We can rewrite this as

$$S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma} \quad = \quad \sum_{\boldsymbol{\pi}} g(s_j) f(x_j) \; - \; \sum_{\boldsymbol{\pi}} g(s_j) f(x_{j-1}) \qquad (\boldsymbol{\beta})$$

Also, we have the identity (it is a collapsing sum)

$$\sum_{\boldsymbol{\pi}} \left(f(x_j) g(x_j) - f(x_{j-1}) g(x_{j-1}) \right) = f(b)g(b) - f(a)g(a). \qquad (\boldsymbol{\gamma})$$

Thus, using Equation $\boldsymbol{\beta}$ and Equation $\boldsymbol{\gamma}$, we have

$$\begin{aligned} f(b)g(b) - f(a)g(a) \; - \; S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \quad &= \quad \sum_{\boldsymbol{\pi}} f(x_j) \left(g(x_j) - g(s_j) \right) \qquad (\boldsymbol{\xi}) \\ &+ \quad \sum_{\boldsymbol{\pi}} f(x_{j-1}) \left(g(s_j) - g(x_{j-1}) \right) \end{aligned}$$

Since $\boldsymbol{\sigma} \subseteq \boldsymbol{\pi}$, we have the ordering

$$a = x_0 \leq s_1 \leq x_1 \leq s_2 \leq x_2 \leq \ldots \leq x_{p-1} \leq s_p \leq x_p = b.$$

Hence, the points above are a refinement of $\boldsymbol{\pi}$ *we will call* $\boldsymbol{\pi'}$. *Relabel the points of* $\boldsymbol{\pi'}$ *as*

$$\boldsymbol{\pi'} = \{y_0, y_1, \ldots, y_q\}$$

and note that the original points of $\boldsymbol{\pi}$ *now form an evaluation set* $\boldsymbol{\sigma'}$ *of* $\boldsymbol{\pi'}$. *We can therefore rewrite Equation* $\boldsymbol{\xi}$ *as*

$$f(b)g(b) - f(a)g(a) \ - \ S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma}) = \sum_{\boldsymbol{\pi}} f(y_j)\Delta g_j \ = \ S(f, g, \boldsymbol{\pi'}, \boldsymbol{\sigma'})$$

Let $I_g = f(b)g(b) - f(a)g(a) \ - \ I_f$. *Then since* $\boldsymbol{\pi_0} \preceq \boldsymbol{\pi} \preceq \boldsymbol{\pi'}$, *we can apply Equation* $\boldsymbol{\alpha}$ *to conclude*

$$\begin{aligned} \epsilon \quad &> \quad \left| S(f, g, \boldsymbol{\pi'}, \boldsymbol{\sigma'} - I_f \right| \\ &= \quad \left| f(b)g(b) - f(a)g(a) \ - \ S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \ - \ I_f \right| \\ &= \quad \left| S(g, f, \boldsymbol{\pi}, \boldsymbol{\sigma}) \ - \ I_g \right| \end{aligned}$$

Since our choice of refinement $\boldsymbol{\pi}$ *of* $\boldsymbol{\pi_0}$ *and evaluation set* $\boldsymbol{\sigma}$ *was arbitrary, we have shown that* $g \in RS[f, a, b]$ *with value*

$$RS(g, f, a, b) = f(x)g(x)\Big|_a^b - RS(f, g, a, b).$$

■

7.2 Step Functions As Integrators

We now turn our attention to the question of what pairs of functions might have a Riemann - Stieljes integral. All we know so far is that if $g(x) = x$ on $[a, b]$ is labeled as $g = \boldsymbol{id}$, then $RS[f, \boldsymbol{id}, a, b] = RI[f, a, b]$.

First, we need to define what we mean by a **Step Function**.

Definition 7.2.1. Step Function

We say $g \in B[a, b]$ *is a Step Function if* g *only has finitely many jump discontinuities on* $[a, b]$ *and* g *is continuous on the intervals between the jump discontinuities. Thus, we may assume there is a non negative integer* p *so that the jump discontinuities are ordered and labeled as*

$$c_0 < c_1 < c_2 < \ldots < c_p$$

and g *is continuous on each subinterval* (c_{k-1}, c_k) *for* $1 \leq k \leq p$.

Comment 7.2.1. *We can see* $g(c_k^-)$ *and* $g(c_k^+)$ *both exist and are finite with* $g(c_k^-)$ *the value* g *has on* (c_{k-1}, c_k) *and* $g(c_k^+)$ *the value* g *has on* (c_k, c_{k+1}). *At the endpoints,* $g(a^+)$ *and* $g(b^-)$ *are also defined. The actual finite values* g *takes on at the points* c_j *are completely arbitrary.*

Lemma 7.2.1. One Jump Step Functions As Integrators
Let $g \in B[a,b]$ be a step function having only one jump at some c in $[a,b]$. Let $f \in B[a,b]$. Then

(i) $f \in C[a,b]$ implies $f \in RS[g,a,b]$ and

- *If $c \in (a,b)$, then $RS(f,g;a,b) = f(c)[g(c^+) - g(c^-)]$.*
- *If $c = a$, then $RS(f,g;a,b) = f(a)[g(a^+) - g(a)]$.*
- *If $c = b$, then $RS(f,g;a,b) = f(b)[g(b) - g(b^-)]$.*

(ii) If $c \in (a,b)$, $f(c^-) = f(c)$ and $g(c^+) = g(c)$, then $f \in RS[g,a,b]$. We can rephrase this as: if c is an interior point, f is continuous from the left at c and g is continuous from the right at c, then $f \in RS[g,a,b]$ and

- *If $c \in (a,b)$, then $RS(f,g;a,b) = f(c)[g(c) - g(c^-)]$.*
- *If $c = a$, then $RS(f,g;a,b) = f(a)[g(a) - g(a)] = 0$.*
- *If $c = b$, then $RS(f,g;a,b) = f(b)[g(b) - g(b^-)]$.*

(iii) If $c \in (a,b)$, $f(c^+) = f(c)$ and $g(c^-) = g(c)$, then $f \in RS[g,a,b]$. We can rephrase this as: if c is an interior point, f is continuous from the right at c and g is continuous from the left at c, then $f \in RS[g,a,b]$ and

- *If $c \in (a,b)$, then $RS(f,g;a,b) = f(c)[g(c^+) - g(c)]$.*
- *If $c = a$, then $RS(f,g;a,b) = f(a)[g(a^+) - g(a)]$.*
- *If $c = b$, then $RS(f,g;a,b) = f(b)[g(b) - g(b)] = 0$.*

Proof. *Let $\boldsymbol{\pi}$ be any partition of $[a,b]$. We will assume that c is a partition point of $\boldsymbol{\pi}$ because if not, we can use the argument we have used before to construct an appropriate refinement as done, for example, in the proof of Lemma 4.5.1. Letting the partition points be*

$$\boldsymbol{\pi} = \{x_0, x_1, \ldots, x_p\},$$

we see there is a partition point $x_{k_0} = c$ with $k_0 \neq 0$ or p. Hence, on $[x_{k_0-1}, x_{k_0}] = [x_{k_0-1}, c]$, $\Delta g_{k_0} = g(c) - g(x_{k_0-1})$. However, since g has a single jump at c, we see that the value $g(x_{k_0-1})$ must be $g(c^-)$. Thus, $\Delta g_{k_0} = g(c) - g(c^-)$. A similar argument shows that $\Delta g_{k_0} = g(c^+) - g(c)$. Further, since g has only one jump, all the other terms Δg_k are zero. Hence, for any evaluation set $\boldsymbol{\sigma}$ in $\boldsymbol{\pi}$, we have

$\boldsymbol{\sigma} = \{s_1, \ldots, s_p\}$ and

$$\begin{aligned} S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma}) &= f(s_{k_0})\Delta g_{k_0} + f(s_{k_0+1}\Delta g_{k_0+1} \\ &= f(s_{k_0})\Bigg(g(c) - g(c^-)\Bigg) + f(s_{k_0+1}\Bigg(g(c^+) - g(c)\Bigg) \\ &= \Bigg(f(s_{k_0}) - f(c) + f(c)\Bigg)\Bigg(g(c) - g(c^-)\Bigg) \\ &+ \Bigg(f(s_{k_0+1} - f(c) + f(c)\Bigg)\Bigg(g(c^+) - g(c)\Bigg) \end{aligned}$$

Thus, we obtain

$$\begin{aligned} S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma}) &= \Bigg(f(s_{k_0}) - f(c)\Bigg)\Bigg(g(c) - g(c^-)\Bigg) \\ &+ \Bigg(f(s_{k_0+1} - f(c)\Bigg)\Bigg(g(c^+) - g(c)\Bigg) \qquad (\boldsymbol{\alpha}) \\ &+ f(c)\Bigg(g(c^+) - g(c^-)\Bigg) \end{aligned}$$

(i)

Subproof. *In this case, f is continuous at c. Let $A = \max\Bigg(\mid g(c) - g(c^-) \mid, \mid g(c^+) - g(c^) \mid\Bigg)$. Then $A > 0$ because g has a jump at c. Since f is continuous at c, given $\epsilon > 0$, there is a $\delta > 0$, so that*

$$\mid f(x) - f(c) \mid \quad < \quad \epsilon/(2A),\ x \in (c - \delta, c + \delta) \cap [a, b]. \qquad (\boldsymbol{\beta})$$

In fact, since c is an interior point of $[a, b]$, we can choose δ so small that $(c - \delta, c + \delta) \subseteq [a, b]$. Now, if $\boldsymbol{\pi_0}$ is any partition with $\| \boldsymbol{\pi_0} \| < \delta$ containing c as a partition point, we can argue as we did in the prefatory remarks to this proof. Thus, there is an index k_0 so that

$$[x_{k_0-1}, x_{k_0} = c] \subseteq (c - \delta, c],\ \ [c = x_{k_0}, x_{k_0+1}] \subseteq [c, c + \delta).$$

This implies that

$$[x_{k_0-1}, x_{k_0+1}] \subseteq (c - \delta, c + \delta)$$

and so the evaluation points, labeled as usual, s_{k_0} and s_{k_0+1} are also in $(c - \delta, c + \delta)$. Applying Equation $\boldsymbol{\beta}$, we have

$$\mid f(s_{k_0}) - f(c)| < \epsilon/(2A),\ \ \mid f(s_{k_0+1}) - f(c)| < \epsilon/(2A.$$

From Equation $\boldsymbol{\alpha}$*, we then have*

$$\begin{aligned}
\left|S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - f(c)\Big(g(c^+) - g(c^-)\Big)\right| &\leq \left|\Big(f(s_{k_0}) - f(c)\Big)\Big(g(c) - g(c^-)\Big)\right| \\
&+ \left|\Big(f(s_{k_0+1} - f(c)\Big)\Big(g(c^+) - g(c)\Big)\right| \\
&< \epsilon/(2A)\Big|g(c) - g(c^-)\Big| + \epsilon/(2A)\Big|g(c^+) - g(c)\Big| \\
&< \epsilon
\end{aligned}$$

Finally, if $\boldsymbol{\pi_0} \preceq \boldsymbol{\pi}$*, then* $\| \boldsymbol{\pi} \| < \delta$ *also and the same argument shows that for any evaluation set* $\boldsymbol{\sigma} \subseteq \boldsymbol{\pi}$*, we have*

$$\left|S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - f(c)\Big(g(c^+) - g(c^-)\Big)\right| < \epsilon$$

This proves that $f \in RS[g,a,b]$ *and* $RS(f,g;a,b) = f(c)\Big(g(c^+) - g(c^-)\Big)$*. Now, if* $c = a$ *or* $c = b$*, the arguments are quite similar, except one sided and we find* $RS(f,g;a,b) = f(a)\Big(g(a^+) - g(a)\Big)$ *or* $RS(f,g;a,b) = f(b)\Big(g(b) - g(b^-)\Big)$*.* □

(ii)

Subproof. *In this case,* f *is continuous from the left at* c *so* $f(c^-) = f(c)$ *and* g *is continuous from the right* $g(c) = g(c^+)$*. Thus, Equation* $\boldsymbol{\alpha}$ *reduces to*

$$S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) = \Big(f(s_{k_0}) - f(c)\Big)\Big(g(c) - g(c^-)\Big) + f(c)\Big(g(c) - g(c^-)\Big) \qquad (\boldsymbol{\alpha'})$$

Let $L = | g(c) - g(c^-) |$*. Then, given* $\epsilon > 0$*, since* f *is continuous from the left, there is a* $\delta > 0$ *so that*

$$| f(x) - f(c) | < \epsilon/L, \ \ x \in (c-\delta, c] \subseteq [a,b].$$

As usual, we can restrict our attention to partitions that contain the point c*. We continue to use* x_i*'s and* s_j*'s to represent points in these partitions and associated evaluation sets. Let* $\boldsymbol{\pi}$ *be such a partition with* $x_{k_0} = c$ *and* $\| \boldsymbol{\pi} \| < \delta$*. Let* $\boldsymbol{\sigma}$ *be any evaluation set of* $\boldsymbol{\pi}$*. Then, we have*

$$[x_{k_0-1}, x_{k_0}] \subseteq (c-\delta, c]$$

and thus

$$| f(s_{k_0}) - f(c) | < \epsilon/L.$$

Hence,

$$\begin{aligned}\left|S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - f(c)\Big(g(c^+ - g(c)\Big)\right| &= \Big|f(s_{k_0} - f(c)\Big|\,\Big|g(c) - g(c^-)\Big| \\ &< \epsilon.\end{aligned}$$

Finally, just as in the previous proof, if $\boldsymbol{\pi_0} \preceq \boldsymbol{\pi}$, then $\|\ \boldsymbol{\pi}\ \| < \delta$ also and the same argument shows that for any evaluation set $\boldsymbol{\sigma} \subseteq \boldsymbol{\pi}$, we have

$$\left|S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - f(c)\Big(g(c) - g(c^-)\Big)\right| < \epsilon$$

This proves that $f \in RS[g,a,b]$ and $RS(f,g;a,b) = f(c)\Big(g(c) - g(c^-)\Big)$. Now, if $c = a$ or $c = b$, the arguments are again similar, except one sided and we find $RS(f,g;a,b) = f(a)\Big(g(a) - g(a)\Big) = 0$ or $RS(f,g;a,b) = f(b)\Big(g(b) - g(b^-)\Big)$. □

(iii)

Subproof. *This is quite similar to the argument presented for Part (ii) above. We find $f \in RS[g,a,b]$ and $RS(f,g;a,b) = f(c)\Big(g(c^+) - g(c)\Big)$. Now, if $c = a$ or $c = b$, the arguments are again similar, except one sided and we find $RS(f,g;a,b) = f(a)\Big(g(a^+) - g(a)\Big) = 0$ or $RS(f,g;a,b) = f(b)\Big(g(b) - g(b)\Big) = 0$.*
□

■

We can then generalize to a finite number of jumps.

Lemma 7.2.2. Finite Jump Step Functions As Integrators
Let g be a step function on $[a,b]$ with jump discontinuities at

$$\{a \leq c_0, c_1, \ldots, c_{k-1}, c_k \leq b\}.$$

Assume $f \in B[a,b]$. Then, if

(i) f is continuous at c_j, or

(ii) f is left continuous at c_j and g is right continuous at c_j, or

(iii) f is right continuous at c_j and g is left continuous at c_j,

then, $f \in RS[f,g,a,b]$ and

$$RS(f,g,a,b) = f(a)\Big(g(a^+) - g(a)\Big) + \sum_{j=0}^{k} f(c_j)\Big(g(c_j^+) - g(c_j^-)\Big) + f(b)\Big(g(b) - g(b^-)\Big).$$

Proof. *Use Lemma 7.2.1 repeatedly.* ■

7.3 Monotone Integrators

The next step is to learn how to deal with integrators that are monotone functions. To do this, we extend the notion of Darboux Upper and Lower Sums in the obvious way.

Definition 7.3.1. Upper and Lower Riemann - Stieljes Darboux Sums

Let $f \in B[a,b]$ and $g \in B[a,b]$ be monotone increasing. Let $\boldsymbol{\pi}$ be any partition of $[a,b]$ with partition points

$$\boldsymbol{\pi} = \{x_0, x_1, \ldots, x_p\}$$

as usual. Define

$$M_j = \sup_{x \in [x_{j-1}, x_j]} f(x), \;\; m_j = \inf_{x \in [x_{j-1}, x_j]} f(x).$$

The Lower Riemann - Stieljes Darboux Sum for f with respect to g on $[a,b]$ for the partition $\boldsymbol{\pi}$ is

$$L(f, g, \boldsymbol{\pi} = \sum_{\boldsymbol{\pi}} m_j \Delta g_j$$

and the Upper Riemann - Stieljes Darboux Sum for f with respect to g on $[a,b]$ for the partition $\boldsymbol{\pi}$ is

$$U(f, g, \boldsymbol{\pi} = \sum_{\boldsymbol{\pi}} M_j \Delta g_j$$

Comment 7.3.1. *It is clear that for any partition $\boldsymbol{\pi}$ and associated evaluation set $\boldsymbol{\sigma}$, that we have the usual inequality chain:*

$$L(f, g, \boldsymbol{\pi}) \;\leq\; S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma} \;\leq\; U(f, g, \boldsymbol{\pi})$$

The following theorems have proofs very similar to the ones we did for Theorem 4.2.1 and Theorem 4.2.2.

Theorem 7.3.1. $\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$ Implies $L(f, g, \boldsymbol{\pi}) \leq L(f, g, \boldsymbol{\pi}')$ and $U(f, g, \boldsymbol{\pi}) \geq U(f, g, \boldsymbol{\pi}')$

Assume g is a bounded monotone increasing function on $[a,b]$ and $f \in B[a,b]$. Then if $\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$, then $L(f, g, \boldsymbol{\pi}) \leq L(f, g, \boldsymbol{\pi}')$ and $U(f, g, \boldsymbol{\pi}) \geq U(f, g, \boldsymbol{\pi}')$.

Theorem 7.3.2. $L(f, g, \boldsymbol{\pi}_1) \leq U(f, g, \boldsymbol{\pi}_2)$

Let $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ be any two partitions in $\boldsymbol{\Pi}[a,b]$. Then $L(f, g, \boldsymbol{\pi}_1) \leq U(f, g, \boldsymbol{\pi}_2)$.

These two theorems allow us to prove the following

Theorem 7.3.3. The Upper And Lower Riemann - Stieljes Darboux Integral Are Finite

Let $f \in B[a,b]$ and let g be a bounded monotone increasing function on $[a,b]$. Let $\mathscr{U} = \{L(f, g, \boldsymbol{\pi}) \mid \boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]\}$ and $\mathscr{V} = \{U(f, g, \boldsymbol{\pi}) \mid \boldsymbol{\pi} \in \boldsymbol{\Pi}[a,b]\}$. Define $L(f,g) = \sup \mathscr{U}$, and $U(f,g) = \inf \mathscr{V}$. Then $L(f,g)$ and $U(f,g)$ are both finite. Moreover, $L(f,g) \;\leq\; U(f,g)$.

We can then define upper and lower Riemann - Stieljes integrals analogous to the way we defined the upper and lower Riemann integrals.

Definition 7.3.2. Upper and Lower Riemann - Stieljes Integrals

Let $f \in B[a,b]$ and g be a bounded, monotone increasing function on $[a,b]$. The Upper and Lower Riemann - Stieljes integrals of f with respect to g are $U(f,g)$ and $L(f,g)$, respectively.

Thus, we can define the Riemann - Stieljes Darboux integral of $f \in B[a,b]$ with respect to the bounded monotone increasing integrator g.

Definition 7.3.3. The Riemann - Stieljes Darboux Integral

Let $f \in B[a,b]$ and g be a bounded, monotone increasing function on $[a,b]$. We say f is Riemann - Stieljes Darboux integrable with respect to the integrator g if $U(f,g) = L(f,g)$. We denote this common value by $RSD(f,g,a,b)$.

7.4 The Riemann - Stieljes Equivalence Theorem

The connection between the Riemann - Stieljes and Riemann - Stieljes Darboux integrals is obtained using an analog of the familiar Riemann Condition we have seen before in Definition 4.2.4.

Definition 7.4.1. The Riemann - Stieljes Criterion For Integrability

Let $f \in B[a,b]$ and g be a bounded monotone increasing function on $[a,b]$. We say the Riemann Condition or Criterion holds for f with respect to g if there is a partition of $[a,b]$, $\boldsymbol{\pi_0}$ so that

$$U(f,g,\boldsymbol{\pi}) - L(f,g,\boldsymbol{\pi}) < \epsilon, \ \boldsymbol{\pi_0} \preceq \boldsymbol{\pi}.$$

We can then prove an equivalence theorem for Riemann - Stieljes and Riemann - Stieljes Darboux integrability.

Theorem 7.4.1. The Riemann Stieljes Integral Equivalence Theorem

Let $f \in B[a,b]$ and g be a bounded monotone increasing function on $[a,b]$. Then the following are equivalent.

(i) $f \in RS[g,a,b]$.

(ii) *Riemann's Criterion holds for f with respect to g.*

(iii) *f is Riemann - Stieljes Darboux Integrable, i.e, $L(f,g) = U(f,g)$, and $RS(f,g;a,b) = RSD(f,g;a,b)$.*

Proof. *The arguments are essentially the same as presented in the proof of Theorem 4.2.4 and hence, you will be asked to go through the original proof and replace occurrences of Δx_j with Δg_j and $b-a$ with $g(b)-g(a)$.* ■

Comment 7.4.1. *We have been very careful to distinguish between Riemann - Stieljes and Riemann - Stieljes Darboux integrability. Since we now know they are equivalent, we can begin to use a common notation. Recall, the common notation for the Riemann integral is $\int_a^b f(x)dx$. We will now begin using the notation $\int_a^b f(x)dg(x)$ to denote the common value $RS(f,g;a,b) = RSD(f,g;a,b)$. We thus know $int_a^b f(x)dx$ is equivalent to the Riemann - Stieljes integral of f with respect to the integrator $g(x) = x$.*

Hence, in this case, we could write $g(x) = \boldsymbol{id}(x) = x$, where $\boldsymbol{id}$ is the identity function. We could then use the notation $\int_a^b f(x)dx = \int_a^b f(x)d\boldsymbol{id}$. However, that is cumbersome. We can easily remember that the identity mapping is simply x itself. So replace $d\boldsymbol{id}$ by $d\boldsymbol{x}$ to obtain $\int_a^b f(x)d\boldsymbol{x}$. The use of the (x) in these notations has always been helpful to allow us to handle substitution type rules, but it is certainly somewhat awkward. A reasonable change of notation would be to go to using boldface for the f and g in these integrals and write $\int_a^b \boldsymbol{f}d\boldsymbol{g}$ giving $\int_a^b \boldsymbol{f}d\boldsymbol{x}$ for the simpler Riemann integral.

You can see no matter what we do the symbolism becomes awkward. For example, suppose $f(x) = sin(x^2)$ on $[0, \pi]$ and $g(x) = x^2$. Then, how do we write $\int_0^\pi \boldsymbol{f}d\boldsymbol{g}$? We will usually abuse our integral notation and write $\int_0^\pi sin(x^2)d(x^2)$.

7.5 Properties Of The Riemann Integral

We can prove the following useful collection of facts about Riemann - Stieljes integrals.

Theorem 7.5.1. Properties Of The Riemann Stieljes Integral

Let the integrator g be bounded and monotone increasing on $[a, b]$. Assume f_1, f_2 and f_3 are in $RS[f, g, a, b]$. Then

(i) $\mid f \mid \in RS[g, a, b]$;

(ii)
$$\left| \int_a^b f(x)dg(x) \right| \leq \int_a^b \mid f \mid dg(x);$$

(iii) $f^+ = \max\{f, 0\} \in RS[g, a, b]$;

(iv) $f^- = \max\{-f, 0\} \in RS[g, a, b]$;

(v)
$$\begin{aligned} \int_a^b f(x)dg(x) &= \int_a^b [f^+(x) - f^-(x)]dg(x) = \int_a^b f^+(x)dg(x) - \int_a^b f^-(x)dg(x) \\ \int_a^b \mid f(x) \mid dg(x) &= \int_a^b [f^+(x) + f^-(x)]dg(x) = \int_a^b f^+(x)dg(x) + \int_a^b f^-(x)dg(x); \end{aligned}$$

(vi) $f^2 \in RS[g, a, b]$;

(vii) $f_1 f_2 \in RS[g, a, b]$;

(viii) If there exists m such that $0 < m \leq f(x)$ for all x in $[a, b]$, then $1/f \in RS[g, a, b]$.

Proof. *The arguments are straightforward modifications of the proof of Theorem 4.3.1 using $b - a = g(b) - g(a)$ and $\Delta x_j = \Delta g_j$.* ∎

We can also easily prove the following fundamental estimate.

Theorem 7.5.2. Fundamental Riemann Stieljes Integral Estimates
Let g be bounded and monotone increasing on $[a,b]$ and let $f \in RS[g,a,b]$. Let $m = \inf_x f(x)$ and let $M = \sup_x f(x)$. Then

$$m(g(b)-g(a)) \leq \int_a^b f(x)dg(x) \leq M(g(b)-g()a).$$

In addition, Riemann - Stieljes integrals are also order preserving as we can modify the proof of Theorem 4.1.3 quite easily.

Theorem 7.5.3. The Riemann Stieljes Integral Is Order Preserving
Let g be bounded and monotone increasing on $[a,b]$ and $f, f_1, f_2 \in RS[g,a,b]$ with $f_1 \leq f_2$ on $[a,b]$. Then the Riemann Stieljes integral is order preserving in the sense that

(i)

$$f \geq 0 \Rightarrow \int_a^b f(x)dg(x) \geq 0;$$

(ii)

$$f_1 \leq f_2 \Rightarrow \int_a^b f_1(x)dg(x) \leq \int_a^b f_2(x)dg(x).$$

We also want to establish the familiar summation property of the Riemann Stieljes integral over an interval $[a,b] = [a,c] \cup [c,b]$. We can modify the proof of the corresponding result in Lemma 4.5.1 as usual to obtain Lemma 7.5.4.

Lemma 7.5.4. The Upper And Lower Riemann - Stieljes Darboux Integral Is Additive On Intervals
Let g be bounded and monotone increasing on $[a,b]$ and $f \in B[a,b]$. Let $c \in (a,b)$. Define

$$\underline{\int_a^b} f(x)\, dg(x) = L(f,g) \text{ and } \overline{\int_a^b} f(x)\, dg(x) = U(f,g)$$

denote the lower and upper Riemann - Stieljes Darboux integrals of f on with respect to g on $[a,b]$, respectively. Then we have

$$\overline{\int_a^b} f(x)dg(x) = \overline{\int_a^c} f(x)dg(x) + \overline{\int_c^b} f(x)dg(x)$$

$$\underline{\int_a^b} f(x)dg(x) = \underline{\int_a^c} f(x)dg(x) + \underline{\int_c^b} f(x)dg(x).$$

Lemma 7.5.4 allows us to prove existence of the Riemann - Stieljes on $[a,b]$ implies it also exists on subintervals of $[a,b]$ and the Riemann - Stieljes value is additive. The proofs are obvious modifications of the proofs of Theorem 4.5.2 and Theorem 4.5.3, respectively.

Theorem 7.5.5. The Riemann Stieljes Integral Exists On Subintervals
Let g be bounded and monotone increasing on $[a,b]$. If $f \in RS[g,a,b]$ and $c \in (a,b)$, then $f \in RS[g,a,c]$ and $f \in RS[g,c,b]$.

Theorem 7.5.6. The Riemann Integral Is Additive On Subintervals
If $f \in RS[g,a,b]$ and $c \in (a,b)$, then

$$\int_a^b f(x)dg(x) = \int_a^c f(x)dg(x) + \int_c^b f(x)dg(x).$$

7.6 Bounded Variation Integrators

We now turn our attention to integrators which are of bounded variation. By Theorem 3.4.3, we know that if $g \in BV[a,b]$, then we can write $g = u - v$ where u and v are monotone increasing on $[a,b]$. Note if h is any other monotone increasing function on $[a,b]$, we could also use the decomposition

$$g = (u+h) - (v+h)$$

as well, so this representation is certainly not unique. We must be very careful when we extend the Riemann - Stieljes integral to bounded variation integrators. For example, even if $f \in RS[g,a,b]$ it does not always follow that $f \in RS[u,a,b]$ and /or $f \in RS[v,a,b]$! However, we can prove that this statement is true if we use a particular decomposition of f. Let $u(x) = V_g(x)$ and $v(x) = V_g(x) - g(x)$ be our decomposition of g. Then, we will be able to show $f \in RS[g,a,b]$ implies $f \in RS[V_g,a,b]$ and $f \in RS[V_g - g,a,b]$.

Theorem 7.6.1. f Riemann Stieljes Integrable With Respect To g Of Bounded Variation Implies Integrable With Respect To V_g and $V_g - g$.
Let $g \in BV[a,b]$ and $f \in RS[g,a,b]$. Then $f \in RS[V_g,a,b]$ and $f \in RS[V_g - g,a,b]$.

Proof. *For convenience of notation, let $u = V_g$ and $v = V_g - g$. First, we show that $f \in RS[u,a,b]$ by showing the Riemann - Stieljes Criterion holds for f with respect to u on $[a,b]$. Fix a positive ϵ. Then there is a partition $\boldsymbol{\pi_0}$ so that*

$$\mid S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - \int_a^b f(x)dg(x) \mid < \epsilon$$

for all refinements $\boldsymbol{\pi}$ of $\boldsymbol{\pi_0}$ and evaluation sets $\boldsymbol{\sigma}$ of $\boldsymbol{\pi}$. Thus, given two such evaluation sets $\boldsymbol{\sigma_1}$ and $\boldsymbol{\sigma_2}$ of a refinement $\boldsymbol{\pi}$, we have

$$\begin{aligned} \mid S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_1}) - S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_2}) \mid &\leq \mid S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_1}) - \int_a^b f(x)dg(x) \mid \\ &+ \mid S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_2}) - \int_a^b f(x)dg(x) \mid \\ &< 2\epsilon. \end{aligned}$$

Hence, we know for $\boldsymbol{\sigma_1} = \{s_1, \ldots, s_p\}$ and $\boldsymbol{\sigma_2} = \{s'_1, \ldots, s'_p\}$, that

$$\mid S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_1} - S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma_2} \mid \ < \ 2\epsilon \qquad (\boldsymbol{\alpha})$$

Now, $u(b) = V_g(b) = \sup_{\boldsymbol{\pi}} \sum_{\boldsymbol{\pi}} \mid \Delta g_j \mid$. Thus, by the Supremum Tolerance Lemma, there is a partition $\boldsymbol{\pi_1}$ so that

$$u(b) - \epsilon < \sum_{\boldsymbol{\pi_1}} \mid \Delta g_j \mid \leq u(b).$$

Then if $\boldsymbol{\pi}$ *refines* $\boldsymbol{\pi_1}$*, we have*

$$u(b) - \epsilon < \sum_{\boldsymbol{\pi_1}} \mid \Delta g_j \mid \leq \sum_{\boldsymbol{\pi}} \mid \Delta g_j \mid \leq u(b).$$

and so for all $\boldsymbol{\pi_1} \preceq \boldsymbol{\pi}$*,*

$$u(b) - \epsilon < \sum_{\boldsymbol{\pi}} \mid \Delta g_j \mid \leq u(b). \qquad (\boldsymbol{\beta})$$

Now let $\boldsymbol{\pi_2} = \boldsymbol{\pi_0} \vee \boldsymbol{\pi_1}$ *and choose any partition* $\boldsymbol{\pi}$ *that refines* $\boldsymbol{\pi_2}$*. Then,*

$$\begin{aligned} \sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \mid \Delta u_j \mid - \mid \Delta g_j \mid \quad &\leq \quad \sum_{\boldsymbol{\pi}} \Big(M_j + m_j \Big) \Delta u_j - \mid \Delta g_j \mid \\ &\leq \quad 2M \sum_{\boldsymbol{\pi}} \mid \Delta u_j \mid - \mid \Delta g_j \mid \end{aligned}$$

where $M = \| f \|_\infty$*. But the term* $\sum_{\boldsymbol{\pi}} \Delta u_j$ *is a collapsing sum which becomes* $u(b) - u(a) = u(b)$ *as* $u(a) = 0$*. We conclude*

$$\sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \mid \Delta u_j \mid - \mid \Delta g_j \mid \quad \leq \quad 2M \left(u(b) - \sum_{\boldsymbol{\pi}} \Big| \Delta g_j \Big| \right.$$

Now by Equation $\boldsymbol{\alpha}$*, for all refinements of* $\boldsymbol{\pi_2}$*, we have*

$$u(b) - \sum_{\boldsymbol{\pi}} \mid \Delta g_j \mid < \epsilon.$$

Hence,

$$\sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \mid \Delta u_j \mid - \mid \Delta g_j \mid \quad \leq \quad 2M\, \epsilon. \qquad (\boldsymbol{\gamma})$$

Next, for any refinement of $\boldsymbol{\pi}$ *of* $\boldsymbol{\pi_2}$*, let the partition points be* $\{x_0, \ldots, x_n\}$ *as usual and define*

$$J^+(\boldsymbol{\pi}) = \{ j \in \boldsymbol{\pi} | \Delta g_j \geq 0 \}, \;\; J^-(\boldsymbol{\pi}) = \{ j \in \boldsymbol{\pi} | \Delta g_j < 0 \}.$$

By the Infimum and Supremum Tolerance Lemma, if $j \in J^+(\boldsymbol{\pi})$*,*

$$\exists s'_j \in [x_{j-1}, x_j] \ni m_j \leq f(s'_j) < m_j + \epsilon/2, \;\; \exists s_j \in [x_{j-1}, x_j] \ni M_j - \epsilon/2 < f(s_j) \leq M_j.$$

It follows

$$f(s_j) - f(s'_j) > M_j - m_j - \epsilon, \;\; j \in J^+(\boldsymbol{\pi}). \qquad (\boldsymbol{\xi})$$

On the other hand, if $j \in J^-(\boldsymbol{\pi})$*, we can find* s_j *and* s'_j *in* $[x_{j-1}, x_j]$ *so that*

$$\exists s'_j \in [x_{j-1}, x_j] \ni m_j \leq f(s_j) < m_j + \epsilon/2, \;\; \exists s_j \in [x_{j-1}, x_j] \ni M_j - \epsilon/2 < f(s'_j) \leq M_j.$$

This leads to

$$f(s'_j) - f(s_j) > M_j - m_j - \epsilon, \;\; j \in J^-(\boldsymbol{\pi}). \tag{$\boldsymbol{\zeta}$}$$

Thus,

$$\begin{aligned}
\sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \mid \Delta g_j \mid \;\; &= \;\; \sum_{j \in J^+(\boldsymbol{\pi})} \Big(M_j - m_j \Big) \Delta g_j \; + \sum_{j \in J^-(\boldsymbol{\pi})} \Big(M_j - m_j \Big) \Big(- \Delta g_j \Big) \\
&< \;\; \sum_{j \in J^+(\boldsymbol{\pi})} \Big(f(s_j) - f(s'_j) \Big) \Delta g_j \; + \; \epsilon \sum_{j \in J^+(\boldsymbol{\pi})} \Delta g_j \\
&+ \;\; \sum_{j \in J^-(\boldsymbol{\pi})} \Big(f(s'_j) - f(s_j) \Big) \Big(- \Delta g_j \Big) \; + \; \epsilon \sum_{j \in J^+(\boldsymbol{\pi})} \Big(- \Delta g_j \Big) \\
&= \;\; \sum_{j \in \boldsymbol{\pi}} \Big(f(s_j) - f(s'_j) \Big) \Delta g_j \; + \; \epsilon \sum_{j \in \boldsymbol{\pi}} \mid \Delta g_j \mid .
\end{aligned}$$

Also, by the definition of the variation function of g, we have

$$\sum_{j \in \boldsymbol{\pi}} \mid \Delta g_j \mid \leq u(b) = V_g(b).$$

Since the points $\{s_1, \ldots, s_n\}$ and $\{s'_1, \ldots, s'_n\}$ are evaluation sets of $\boldsymbol{\pi}$, we can apply Equation $\boldsymbol{\alpha}$ to conclude

$$\begin{aligned}
\mid S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma_1} - S(f, g, \boldsymbol{\pi}, \boldsymbol{\sigma_2} \mid \;\; &= \;\; \sum_{j \in \boldsymbol{\pi}} \Big(f(s_j) - f(s'_j) \Big) \Delta g_j \\
&< \;\; 2\epsilon.
\end{aligned}$$

Hence,

$$\sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \mid \Delta g_j \mid \;\; < \;\; 2\epsilon + \epsilon u(b) \;\; = \;\; (2 + u(b))\epsilon. \tag{$\boldsymbol{\theta}$}$$

Then, using Equation $\boldsymbol{\gamma}$ and Equation $\boldsymbol{\theta}$, we find

$$\begin{aligned}
\sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \Delta u_j \;\; &= \;\; \sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \Big(\Delta u_j - \mid \Delta g_j \mid \Big) \; + \sum_{\boldsymbol{\pi}} \Big(M_j - m_j \Big) \Big(\mid \Delta g_j \mid \Big) \\
&< \;\; 2M\epsilon \; + \; (2 + u(b))\epsilon \; = \; (2M + 2 + u(b))\epsilon.
\end{aligned}$$

Letting $A = 2M + 2 + u(b)$, and recalling that $u = V_g$, we have

$$U(f, V_g, \boldsymbol{\pi}) - L(f, V_g, \boldsymbol{\pi}) < A\epsilon$$

for any refinement $\boldsymbol{\pi}$ of $\boldsymbol{\pi_2}$. Hence, f satisfies the Riemann - Stieljes Criterion with respect to V_g on $[a, b]$. We conclude $f \in RS[V_g, a, b]$.

Thus, $f \in RS[g, a, b]$ and $f \in RS[V_g, a, b]$ and by Theorem 7.1.1, we have $f \in RS[V_g - g, a, b]$ also.

■

Theorem 7.6.2. Products And Reciprocals Of Functions Riemann Stieljes Integrable With Respect To g Of Bounded Variation Are Also Integrable

Let $g \in BV[a,b]$ and $f, f_1, f_2 \in RS[g,a,b]$. Then

(i) $f^2 \in RS[g,a,b]$

(ii) $f_1 f_2 \in RS[g,a,b]$

(iii) If there is a positive constant m, so that $|f(x)| > m$ for all x in $[a,b]$, then $1/f \in Rs[g,a,b]$.

Proof. *(i)*

Subproof. *Since $f \in RS[g,a,b]$, $f \in RS[V_g,a,b]$ and $f \in RS[V_g - g,a,b]$ by Theorem 7.6.1. Hence, by Theorem 7.5.1, $f^2 \in RS[V_g,a,b]$ and $f^2 \in RS[V_g - g,a,b]$. Then, by the linearity of the Riemann Stieljes integral for monotone integrators, Theorem 7.1.1, we have $f^2 \in RS[V_g - (V_g - g) = g,a,b]$.* □

(ii)

Subproof. *$f_1, f_2 \in RS[g,a,b]$ implies $f_1, f_2 \in RS[V_g,a,b]$ and $f_1, f_2 \in RS[V_g - g,a,b]$. Thus, using reasoning just like that in Part (i), we have $f_1 f_2 \in RS[g,a,b]$.* □

(iii)

Subproof. *By our assumptions, we know $1/f \in RS[V_g,a,b]$ and $1/f \in RS[V_g - g,a,b]$. Thus, by the linearity of the Riemann Stieljes integral with respect to monotone integrators, $1/f \in RS[g,a,b]$.* □

■

Theorem 7.6.3. The Riemann Stieljes Integral Is Additive On Subintervals

Let $g \in BV[a,b]$ and $f \in RS[g,a,b]$. Then, if $a \le c \le b$,

$$\int_a^b f(x)dg(x) = \int_a^c f(x)dg(x) + \int_c^b f(x)dg(x).$$

Proof. *From Theorem 7.5.6, we know*

$$\begin{aligned}\int_a^b f(x)dV_g(x) &= \int_a^c f(x)dV_g(x) + \int_c^b f(x)dV_g(x)\\ \int_a^b f(x)d(V_g - g)(x) &= \int_a^c f(x)d(V_g - g)(x) + \int_c^b f(x)d(V_g - g)(x).\end{aligned}$$

Also, we know

$$\int_a^b f(x)dg(x) = \int_a^b f(x)dV_g(x) - \int_c^b f(x)d(V_g - g)(x)$$

and so the result follows. ■

Chapter 8

Further Riemann Stieljes Results

We know quite a bit about the Riemann Stieljes integral in theory. However, we do not know how to compute a Riemann Stieljes integral and we only know that Riemann Stieljes integrals exist for a few type of integrators: those that are bounded with a finite number of jumps and the identity integrator $g(x) = x$. It is time to learn more.

8.1 The Riemann - Stieljes Fundamental Theorem Of Calculus

As you might expect, we can prove a Riemann - Stieljes variant of the Fundamental Theorem Of Calculus.

Theorem 8.1.1. Riemann Stieljes Fundamental Theorem Of Calculus
Let $g \in BV[a,b]m$ $f \in RS[g,a,b]$. Define $F : [a,b] \to \Re$ by

$$F(x) = \int_a^x f(t) dg(t).$$

Then

(i) $F \in BV[a,b]$,

(ii) If g is continuous at c in $[a,b]$, then F is continuous at c.

(iii) If g is monotone and if at c is in $[a,b]$, $g'(c)$ exists and f is continuous at c, then $F'(c)$ exists with

$$F'(c) = f(c)\, g'(c).$$

Proof. *First, assume g is monotone increasing and $g(a) < g(b)$. Let $\boldsymbol{\pi}$ be a partition of $[a,b]$. Then, we immediately have the fundamental estimates*

$$m(g(b) - g(a)) \leq L(f,g) \leq U(f,g) \leq M(g(b) - g(a)),$$

where m and M are the infimum and supremum of f on $[a,b]$ respectively. Since $f \in RS[g,a,b]$, we then have

$$m(g(b)-g(a)) \leq \int_a^b f dg \leq M(g(b)-g(a)).$$

or

$$m \leq \frac{\int_a^b f dg}{g(b)-g(a)} \leq M.$$

Let $K(a,b) = \int_a^b f dg/(g(b)-g(a))$. Then, $m \leq K(a,b) \leq M$ and $\int_a^b f dg = K(a,b)(g(b)-g(a))$.

Now assume $x < y$ in $[a,b]$. Since $f \in RS[g,a,b]$, by Theorem 7.5.5, $f \in RS[g,x,y]$. By the argument just presented, we can show there is a number $K(x,y)$ so that

$$\begin{aligned} K(x,y) &= \int_x^y f dg/(g(y)-g(x)), \\ m \leq \inf_{t\in[x,y]} f(t) &\leq K(x,y) \leq \sup_{t\in[x,y]} f(t) \leq M \\ \int_x^y f dg &= K(x,y)(g(y)-g(x)) \end{aligned} \qquad (\boldsymbol{\alpha})$$

(i)

Subproof. *We show $f \in BV[a,b]$. Let $\boldsymbol{\pi}$ be a partition of $[a,b]$. Then, labeling the partition points in the usual way,*

$$\begin{aligned} \sum_{\boldsymbol{\pi}} \mid \Delta F_j \mid &= \sum_{\boldsymbol{\pi}} \mid \Delta F(x_j) - F(x_{j-1} \mid \\ &= \sum_{\boldsymbol{\pi}} \mid \int_{x_{j-1}}^{x_j} f dg \mid \\ &= \sum_{\boldsymbol{\pi}} \mid K(x_{j-1},x_j) \mid\mid g(x_j) - g(x_{j-1}) \mid = \sum_{\boldsymbol{\pi}} \mid K(x_{j-1},x_j) \mid\mid \Delta g_j \mid \end{aligned}$$

using Equation $\boldsymbol{\alpha}$ on each subinterval $[x_{j-1},x_j]$. However, we know each $m \leq K(x_{j-1},x_j) \leq M$ and so

$$\begin{aligned} \sum_{\boldsymbol{\pi}} \mid \Delta F_j \mid &\leq \parallel f \parallel_\infty \sum_{\boldsymbol{\pi}} \mid \Delta g_j \mid \\ &= \parallel f \parallel_\infty (g(b)-g(a)), \end{aligned}$$

as g is monotone increasing. Since this inequality holds for all partitions of $[a,b]$, we see

$$V(F;a,b) \leq \parallel f \parallel_\infty (g(b)-g(a))$$

implying $F \in BV[a,b]$. □

(ii)

Subproof. *Let g be continuous at c. Then given a positive ϵ, there is a $\delta > 0$, so that*

$$\mid g(c) - g(y) \mid \quad < \quad \epsilon/(1+ \parallel f \parallel_\infty), \ \mid y - c \mid < \delta, \ y \in [a,b].$$

For any such y, apply Equation $\boldsymbol{\alpha}$ *to the interval $[c, y]$ or $[y, c]$ depending on whether $y > c$ or vice - versa. For concreteness, let's look at the case $y > c$. Then, there is a $K(c, y)$ so that $m \leq K(c, y) \leq M$ and $\int_c^y f(t)dg(t) = K(c, y)(g(y) - g(c))$. Thus, since y is within δ of c, we have*

$$\left| \int_c^y f(t)dg(t) \right| = \mid K(c, y) \mid \mid g(y) - g(c) \mid \leq \parallel f \parallel_\infty \; \epsilon/(1 + \parallel f \parallel_\infty) \; < \; \epsilon.$$

We conclude that if $y \in [c, c + \delta)$, then $\left| \int_c^y f(t)dg(t) \right| < \epsilon$. A similar argument holds for $y \in (c - \delta, c]$. Combining, we see $y \in (c - \delta, c + \delta)$ and in $[a, b]$ implies

$$\left| F(y) - F(c) \right| = \left| \int_c^y f(t)dg(t) \right| \; < \; \epsilon.$$

So F is continuous at c. □

(iii)

Subproof. *If $c \in [a, b]$, $g'(c)$ exists and f is continuous at c, we must show that $F'(c) = f(c)g'(c)$. Let a positive ϵ be given. Then,*

$$\exists \delta_1 \ni \left| \frac{g(y) - g(c)}{y - c} - g'(c) \right| \quad < \quad \epsilon, \;\; 0 < \mid y - c \mid < \delta_1, \;\; y \in [a, b]. \qquad (\boldsymbol{\beta})$$

and

$$\exists \delta_2 \ni \Big| f(y) - f(c) \mid \quad < \quad \epsilon, \;\; \mid y - c \mid < \delta_2, \;\; y \in [a, b]. \qquad (\boldsymbol{\gamma})$$

Choose any $\delta < \min(\delta_1, \delta_2)$. Let y be in $\Big((c - \delta, c) \cup (c, c + \delta) \Big) \cap [a, b]$. We are interested in the interval I with endpoints c and y which is either of the form $[c, y]$ or vice - versa. Apply Equation $\boldsymbol{\alpha}$ *to this interval. We find there is a $K(I)$ that satisfies*

$$\inf_{t \in I} f(t) \leq K(I) \leq \sup_{t \in I} f(t)$$

and

$$\int_c^y f(t)dg(t) = K([c, y])(g(y) - g(c)), \;\; y > c$$

or

$$\int_y^c f(t)dg(t) \quad = \quad K([y, c])(g(c) - g(y)), \;\; y < c$$

or

$$-\int_c^y f(t)dg(t) \quad = \quad K([y, c])(g(c) - g(y)), \;\; y < c$$

which gives

$$\int_c^y f(t)dg(t) = K([y,c])(g(y)-g(c)), \;\; y < c.$$

So we conclude we can write

$$\int_c^y f(t)dg(t) = K(I)(g(y)-g(c)).$$

where $K(I)$ *denotes* $K([c,y])$ *or* $K([y,c])$ *depending on where* y *is relative to* c. *Next, since* $\delta < \min(\delta_1, \delta_2)$, *both Equation* $\boldsymbol{\alpha}$ *and Equation* $\boldsymbol{\beta}$ *holds. Thus,*

$$f(c)-\epsilon < f(t) < f(c)+\epsilon, \;\; y \in \Big((c-\delta,c)\cup(c,c+\delta)\Big)\cap[a,b].$$

This tells us that $\sup_{t\in I} f(t) \le f(c)+\epsilon$ *and* $\inf_{t\in I} f(t) \ge f(c)-\epsilon$. *Thus,*

$$f(c)-\epsilon \le K([c,y]), K([y,c]) \le f(c)+\epsilon$$

or $\mid K([c,y]) - f(c) \mid < \epsilon$ *and* $\mid K([y,c]) - f(c) \mid < \epsilon$. *Finally, consider*

$$\begin{aligned}
\left|\frac{F(y)-F(c)}{y-c} - f(c)g'(c)\right| &= \left|\frac{K(I)(g(y)-g(c)}{y-c} - f(c)g'(c)\right| \\
&= \left|\frac{K(I)(g(y)-g(c)}{y-c} - f(c)g'(c) + K(I)g'(c) - K(I)g'(c)\right| \\
&\le \mid K(I) \mid \left|\frac{(g(y)-g(c)}{y-c} - g'(c)\right| + \left|K(I)-f(c)\right| \mid g'(c) \mid \\
&< \parallel f \parallel_\infty \epsilon + \mid g'(c) \mid \epsilon.
\end{aligned}$$

Since ϵ *is arbitrary, this shows* F *is differentiable at* c *with value* $f(c)g'(c)$. □

This proves the proposition for the case that g *is monotone. To finish the proof, we note if* $g \in BV[a,b]$, *then* $g = V_g - (V_g - g)$ *is the standard decomposition of* g *into the difference of two monotone increasing functions. Let* $F_1(x) = \int_a^x f(t)d(V_g)(t)$ *and* $F_2(x) = \int_a^x f(t)d(V_g-g)(t)$. *From Part (i), we see* $F = F_1 - F_2$ *is of bounded variation. Next, if* g *is continuous at* c, *so is* V_g *and* $V_g - g$ *by Theorem 3.5.3. So by Part (ii),* F_1 *and* F_2 *are continuous at* c. *This implies* F *is continuous at* c. ■

8.2 Existence Results

We begin by looking at continuous integrands.

Theorem 8.2.1. Integrand Continuous and Integrator Of Bounded Variation Implies Riemann - Stieljes Integral Exists

If $f \in C[a,b]$ *and* $g \in BV[a,b]$, *then* $f \in RS[g,a,b]$.

Proof. *Let's begin by assuming* g *is monotone increasing. We may assume without loss of generality that* $g(a) < g(b)$. *Let* $K = g(b) - g(a) > 0$. *Since* f *is continuous on* $[a,b]$, f *is uniformly continuous on*

$[a,b]$. *Hence, given a positive* ϵ, *there is a positive* δ *so that*

$$\mid f(s) - f(t) \mid < \epsilon/K, \ \mid t - s \mid < \delta, \ t, s \in [a,b].$$

Now, repeat the proof of Theorem 4.4.1 which shows that if f *is continuous on* $[a,b]$, *then* $f \in RI[a,b]$, *but replace all the* Δx_j *by* Δg_j. *This shows that* f *satisfies the Riemann - Stieljes Criterion for integrability. Thus, by the equivalence theorem,* $f \in RS[g,a,b]$.

Next, let $g \in BV[a,b]$. *Then* $g = V_g - (V_g - g)$ *as usual. Since* V_g *and* $V_g - g$ *are monotone increasing, we can apply our first argument to conclude* $f \in RS[V_g,a,b]$ *and* $f \in RS[V_g - g,a,b]$. *Then, by the linearity of the Riemann - Stieljes integral with respect to the integrator, Theorem 7.1.1, we have* $f \in RS[g,a,b]$ *with*

$$\int_a^b f dg = \int_a^b f dv_g - \int_a^b f d(V_g - g).$$

■

Next, we let the integrand be of bounded variation.

Theorem 8.2.2. Integrand Bounded Variation and Integrator Continuous Implies Riemann - Stieljes Integral Exists

If $f \in BV[a,b]$ *and* $g \in C[a,b]$, *then* $f \in RS[g,a,b]$.

Proof. *If* $f \in BV[a,b]$ *and* $g \in C[a,b]]$, *then by the previous theorem, Theorem 8.2.1,* $g \in RS[f,a,b]$. *Now apply integration by parts, Theorem 7.1.2, to conclude* $f \in RS[g,a,b]$. ■

What if the integrator is differentiable?

Theorem 8.2.3. Integrand Continuous and Integrator Continuously Differentiable Implies Riemann - Stieljes Integrable

Let $f \in C[a,b]$ *and* $g \in C^1[a,b]$. *Then* $f \in RS[g,a,b]$, $fg' \in RI[a,b]$ *and*

$$\int_a^b f(x)dg(x) = \int_a^b f(x)g'(x)dx$$

where the integral on the left side is a traditional Riemann integral.

Proof. *Pick an arbitrary positive* ϵ. *Since* g' *is continuous on* $[a,b]$, g' *is uniformly continuous on* $[a,b]$. *Thus, there is a positive* δ *so that*

$$\mid g'(s) - g'(t) \mid \quad < \quad \epsilon, \ \mid s - t \mid < \delta, \ s, t \in [a,b]. \qquad (\boldsymbol{\alpha})$$

Since g' *is continuous on* $[a,b]$, *there is a number* M *so that* $\mid g(x) \mid \leq M$ *for all* x *in* $[a,b]$. *We conclude that* $g \in BV[a,b]$ *by Theorem 3.3.3. Now apply Theorem 8.2.1, to conclude* $f \in RS[g,a,b]$. *Thus, there is a partition* $\boldsymbol{\pi_0}$ *of* $[a,b]$, *so that*

$$\left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - \int_a^b f dg \right| \quad < \quad \epsilon, \ \boldsymbol{\pi_0} \preceq \boldsymbol{\pi}, \ \boldsymbol{\sigma} \subseteq \boldsymbol{\pi}. \qquad (\boldsymbol{\beta})$$

Further, since fg' is continuous on $[a,b]$, $fg' \in RI[a,b]$ and so $\int_a^b fg'$ exists also.

Now let $\boldsymbol{\pi_1}$ be a refinement of $\boldsymbol{\pi_0}$ with $\|\, \boldsymbol{\pi_1} \,\| < \delta$. Then we can apply Equation $\boldsymbol{\beta}$ to conclude

$$\left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - \int_a^b f dg \right| \quad < \quad \epsilon, \ \boldsymbol{\pi_1} \preceq \boldsymbol{\pi}, \ \boldsymbol{\sigma} \subseteq \boldsymbol{\pi}. \qquad (\gamma)$$

Next, apply the Mean Value Theorem to g on the subintervals $[x_{j-1}, x_j]$ from partition $\boldsymbol{\pi}$ for which Equation $\boldsymbol{\gamma}$ holds. Then, $\Delta g_j = g'(t_j)(x_j - x_{j-1})$ for some t_j in (x_{j-1}, x_j). Hence,

$$S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) = \sum_{\boldsymbol{\pi}} f(s_j)\Delta g_j = \sum_{\boldsymbol{\pi}} f(s_j) g'(t_j)\Delta x_j.$$

Also, we see

$$S(fg',\boldsymbol{\pi},\boldsymbol{\sigma}) = \sum_{\boldsymbol{\pi}} f(s_j) g'(s_j)\Delta x_j.$$

Thus, we can compute

$$\begin{aligned} \left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - S(fg',\boldsymbol{\pi},\boldsymbol{\sigma}) \right| &= \left| \sum_{\boldsymbol{\pi}} f(s_j) \Big(g'(t_j) - g'(s_j)\Delta x_j \Big) \right. \\ &\le \|\, f \,\|_\infty \sum_{\boldsymbol{\pi}} \left| g'(t_j) - g'(s_j)\Delta x_j \right|. \end{aligned}$$

By Equation $\boldsymbol{\alpha}$, since $\|\, \boldsymbol{\pi} \,\| < \delta$, $|t_j - s_j| < \delta$ and so $|g'(t_j) - g'(s_j)| < \epsilon$. We conclude

$$\left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - S(fg',\boldsymbol{\pi},\boldsymbol{\sigma}) \right| \quad < \quad \epsilon \,\|\, f \,\|_\infty \sum_{\boldsymbol{\pi}} \Delta x_j \ = \ \epsilon \,\|\, f \,\|_\infty \, (b-a). \qquad (\boldsymbol{\xi})$$

Thus,

$$\begin{aligned} \left| S(fg',\boldsymbol{\pi},\boldsymbol{\sigma}) - \int_a^b f dg \right| &\le \left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - S(fg',\boldsymbol{\pi},\boldsymbol{\sigma}) \right| + \left| S(f,g,\boldsymbol{\pi},\boldsymbol{\sigma}) - \int_a^b f dg \right| \\ &< \epsilon \,\|\, f \,\|_\infty \, (b-a) + \epsilon \end{aligned}$$

by Equation $\boldsymbol{\gamma}$ and Equation $\boldsymbol{\xi}$. This proves the desired result. ■

It should be easy to see that the assumptions of Theorem 8.2.3 can be relaxed. Consider

Theorem 8.2.4. Integrand Riemann Integrable and Integrator Continuously Differentiable Implies Riemann - Stieljes Integrable

Let $f \in C[a,b]$ and $g \in C^1[a,b]$. Then $f \in RS[g,a,b]$, $fg' \in RI[a,b]$ and

$$\int_a^b f(x)dg(x) = \int_a^b f(x)g'(x)dx$$

where the integral on the left side is a traditional Riemann integral.

Proof. *We never use the continuity of f in the proof given for Theorem 8.2.3. All we use is the fact that f is Riemann integrable. Hence, we can use the proof of Theorem 8.2.3 without change to find*

$$\left| S(fg', \boldsymbol{\pi}, \boldsymbol{\sigma}) - \int_a^b f dg \right| \quad < \quad \epsilon \mid\mid f \mid\mid_\infty (b-a) + \epsilon.$$

This tells us that fg' is Riemann integrable on $[a,b]$ with value $\int_a^b f dg$. ■

8.3 Worked Out Examples Of Riemann Stieljes Computations

How *do* we compute a Riemann Stieljes integral? Let's look at some example.

Example 8.3.1. *Let f and g be defined on $[0,2]$ by*

$$f(x) \quad = \quad \begin{cases} x, & x \in Q \cap [0,2] \\ 2-x, & x \in Ir \cap [0,2], \end{cases} \quad g(x) \ = \ \begin{cases} 1, & 0 \leq x < 1 \\ 3, & 1 \leq x \leq 2. \end{cases}$$

Does $\int f dg$ exist?

Solution 8.3.1. *We can answer this two ways so far. Method 1: We note f is continuous at 1 (you should be able to do a traditional $\epsilon - \delta$ proof of this fact!) and since g has a jump at 1, we can look at Lemma 7.2.1 to see that f is indeed Riemann - Stieljes with respect to g. The value is given by*

$$\int_0^2 f dg = f(1)(g(1^+) - g(1^-)) = 1(3-1) = 2.$$

Method 2: We can compute the integral using a partition approach. Let $\boldsymbol{\pi}$ be a partition of $[0,2]$. We may assume without loss of generality that $1 \in \boldsymbol{\pi}$ (recall all of our earlier arguments that allow us to make this statement!). Hence, there is an index k_0 such that $x_{k_0} = 1$. We have

$$L(f,g,\boldsymbol{\pi}) = \left(\inf_{x \in [x_{k_0-1}, 1]} f(x) \right) \left(g(1) - g(x_{k_0-1}) \right) + \left(\inf_{x \in [1, x_{k_0+1}]} f(x) \right) \left(g(x_{k_0+1}) - g(1) \right).$$

Now use how g is defined to see,

$$L(f,g,\boldsymbol{\pi}) = \left(\inf_{x \in [x_{k_0-1}, 1]} f(x) \right) \left(3 - 1 \right) + \left(\inf_{x \in [1, x_{k_0+1}]} f(x) \right) \left(3 - 3 \right).$$

Hence,

$$L(f,g,\boldsymbol{\pi}) = 2 \left(\inf_{x \in [x_{k_0-1}, 1]} f(x) \right).$$

If you graphed x and $2-x$ simultaneously on $[0,2]$, you would see that they cross at 1 and x is below $2-x$ before 1. This graph works well for f even though we can only use the graph of x when x is rational and the graph of $2-x$ when x is irrational. We can see *in our mind how to do the visualization. For this mental picture, you should be able to see that the infimum of f on $[x_{k_0-1}, 1]$ will be the value x_{k_0-1}. We have thus found that $L(f,g,\boldsymbol{\pi}) = 2x_{k_0-1}$. A similar argument will show that $U(f,g,\boldsymbol{\pi}) = 2(2 - x_{k_0-1})$. This immediately implies that $L(f,g) = U(f,g) = 2$.*

Example 8.3.2. *Let f be any bounded function which is discontinuous from the left at 1 on $[0,2]$. Again, let g be defined on $[0,2]$ by*

$$g(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 3, & 1 \leq x \leq 2. \end{cases}$$

Does $\int f dg$ exist?

Solution 8.3.2. *First, since we know f is not continuous from the left at 1 and g is continuous from the right at 1, the conditions of Lemma 7.2.1 do not hold. So it is possible this integral does not exist. We will in fact show this using arguments that are similar to the previous example. Again, $\boldsymbol{\pi}$ is a partition which has $x_{k_0} = 1$. We find*

$$L(f,g,\boldsymbol{\pi}) = \left(\inf_{x \in [x_{k_0-1},1]} f(x)\right)\left(3-1\right), \; U(f,g,\boldsymbol{\pi}) = \left(\sup_{x \in [x_{k_0-1},1]} f(x)\right)\left(3-1\right).$$

Since we can choose $x_{k_0} - 1$ as close to 1 as we wish, we see

$$\inf_{x \in [x_{k_0-1},1]} f(x) \rightarrow \min(f(1^-), f(1))$$
$$\sup_{x \in [x_{k_0-1},1]} f(x) \rightarrow \max(f(1^-), f(1))$$

But f is discontinuous from the left at 1 and so $f(1^-) \neq f(1)$. For concreteness, let's assume $f(1^-) < f(1)$ (the argument the other way is very similar). We see $L(f,g) = 2f(1^-)$ and $U(f,g) = 2f(1)$. Since these values are not the same, f is not Riemann Stieljes integrable with respect to g by the Riemann - Stieljes equivalence theorem, Theorem 7.4.1.

Example 8.3.3. *Let f be any bounded function which is continuous from the left at 1 on $[0,2]$. Again, let g be defined on $[0,2]$ by*

$$g(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 3, & 1 \leq x \leq 2. \end{cases}$$

Does $\int f dg$ exist?

Solution 8.3.3. *First, since we know f is continuous from the left at 1 and g is continuous from the right at 1, the conditions of Lemma 7.2.1 do hold. So this integral does exist. Using Lemma 7.2.1, we see*

$$\int_0^2 f dg = f(1)(g(1^+) - g(1^-)) = 2f(1).$$

We can also show this using partition arguments as we have done before. Again, $\boldsymbol{\pi}$ is a partition which has $x_{k_0} = 1$. Again, we have

$$L(f,g,\boldsymbol{\pi}) = \left(\inf_{x \in [x_{k_0-1},1]} f(x)\right)\left(3-1\right), \; U(f,g,\boldsymbol{\pi}) = \left(\sup_{x \in [x_{k_0-1},1]} f(x)\right)\left(3-1\right).$$

Since we can choose $x_{k_0} - 1$ *as close to* 1 *as we wish, we see*

$$\inf_{x \in [x_{k_0-1},1]} f(x) \to \min(f(1^-), f(1))$$
$$\sup_{x \in [x_{k_0-1},1]} f(x) \to \max(f(1^-), f(1))$$

But f *is continuous from the left at* 1, $f(1^-) = f(1)$. *We see* $L(f,g) = 2f(1)$ *and* $U(f,g) = 2f(1)$. *Since these values are the same,* f *is Riemann Stieljes integrable with respect to* g *by the Riemann - Stieljes equivalence theorem, Theorem 7.4.1.*

Example 8.3.4. *Define a step function* g *on* $[0, 12]$ *by*

$$g(x) = \begin{cases} 0, & 0 \le x < 2 \\ \sum_{j=2}^{\lfloor x \rfloor} (j-1)/36, & 2 \le x < 8 \\ 21/36 + \sum_{j=8}^{\lfloor x \rfloor} (13-j)/36, & 8 \le x \le 12 \end{cases}$$

where $\lfloor x \rfloor$ *is the greatest integer which is less than or equal to* x. *The function* g *is everywhere continuous from the right and represents the probability of rolling a number* $j \le x$. *It is called the* cumulative probability distribution function of a fair pair of dice. *The Riemann - Stieljes integral* $\mu = \int_0^{12} x dg(x)$ *is called the* mean *of this distribution. The* variance *of this distribution is denoted by* $\boldsymbol{\sigma^2}$ *(unfortunate choice, isn't it as that is the letter we use to denote evaluation sets of partitions!) and defined to be*

$$\boldsymbol{\sigma^2} = \int_0^{12} (x-\mu)^2 dg(x).$$

Compute μ *and* $\boldsymbol{\sigma^2}$.

Solution 8.3.4. *Since* $f(x) = x$ *is continuous on* $[0, 12]$, *Lemma 7.2.2 applies and we have*

$$\int_0^{12} x dg(x) = \sum_{j=2}^{12} j \Big(g(j+) - g(j^-) \Big).$$

The evaluations are a bit messy.

$$\begin{aligned}
36(g(2+) - g(2^-)) &= 36(g(2) - g(2^-)) \;=\; 1 - 0 = 1\\
36(g(3+) - g(3^-)) &= 36(g(3) - g(3^-)) \;=\; 3 - 1 = 2\\
36(g(4+) - g(4^-)) &= 36(g(4) - g(4^-)) \;=\; 6 - 3 = 3\\
36(g(5+) - g(5^-)) &= 36(g(5) - g(5^-)) \;=\; 10 - 6 = 4\\
36(g(6+) - g(6^-)) &= 36(g(6) - g(6^-)) \;=\; 15 - 10 = 5\\
36(g(7+) - g(7^-)) &= 36(g(7) - g(7^-)) \;=\; 21 - 15 = 6\\
36(g(8+) - g(8^-)) &= 36(g(8) - g(8^-)) \;=\; 26 - 21 = 5\\
36(g(9+) - g(9^-)) &= 36(g(9) - g(9^-)) \;=\; 30 - 26 = 4\\
36(g(10+) - g(10^-)) &= 36(g(10) - g(10^-)) \;=\; 33 - 30 = 3\\
36(g(11+) - g(11^-)) &= 36(g(11) - g(11^-)) \;=\; 35 - 33 = 2\\
36(g(12+) - g(12^-)) &= 36(g(12) - g(12^-)) \;=\; 36 - 35 = 1
\end{aligned}$$

Thus,

$$\begin{aligned}
\int_0^{12} x dg(x) &= \Big(2(1) + 3(2) + 4(3) + 5(4) + 6(5) + 7(6)\\
&\quad +8(5) + 9(4) + 10(3) + 11(2) + 12(1) \Big)/36\\
&= \Big(2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12 \Big)/36\\
&= 252/36 \;=\; 7.
\end{aligned}$$

So, the mean *or* expected value *of a single roll of a fair pair of dice is* 7*. To find the* variance*, we calculate*

$$\begin{aligned}
\boldsymbol{\sigma^2} &= \int_0^{12} (x-7) dg(x)\\
&= \sum_{j=2}^{12} (j-7)^2 \Big(g(j+) - g(j^-) \Big)\\
&= \Big(25(1) + 16(2) + 9(3) + 4(4) + 1(5) + 0(6) + 1(5) + 4(4) + 9(3) + 16(2) + 25(1) \Big)/36\\
&= \Big(25 + 32 + 27 + 16 + 5 + 5 + 16 + 27 + 32 + 25 \Big)/36\\
&= 210/36 \;=\; 35/6.
\end{aligned}$$

Example 8.3.5. *Let $f(x) = e^x$ and let g be defined on $[0, 2]$ by*

$$g(x) = \begin{cases} x^2, & 0 \le x \le 1\\ x^2 + 1, & 1 < x \le 2. \end{cases}$$

Show $\int f dg$ exists and evaluate it.

Solution 8.3.5. *Since g is monotone, $\int_0^2 f dg$ exists. We can thus decompose g into its continuous and saltus part. We find*

$$g_c(x) = x^2, \; s_g(x) = \begin{cases} 0, & 0 \le x \le 1 \\ 1, & 1 < x \le 2. \end{cases}$$

The saltus integral is evaluated using Lemma 7.2.1. The integrand is continuous and the jump is at 1, so we have

$$\begin{aligned} \int_0^2 f\, ds_g &= \int_0^2 e^x\, ds_g(x) \\ &= e^1(s_g(1^+) - s_g(1^-)) = e(1-0) = e. \end{aligned}$$

and for the continuous part, we can use the fact the integrator is continuously differentiable on $[0,2]$ to apply Theorem 8.2.3 to obtain

$$\int_0^2 f\, dg_c = \int_0^2 e^x\, d(x^2) = \int_0^2 e^x\, 2x dx = 2(e^2+1)$$

Thus,

$$\begin{aligned} \int_0^2 f dg &= \int_0^2 f\, dg_c + \int_0^2 f\, ds_g \\ &= 2(e^2+1) + e. \end{aligned}$$

We can also do this by integration by parts, Theorem 7.1.2. Since $f \in RS[g,0,2]$, it follows that $g \in RS[f,0,2]$ and

$$\begin{aligned} \int_0^2 f(x)dg(x) &= e^x\, g(x)\Big|_0^2 - \int_0^2 g(x)df(x) \\ &= e^2 g(2) - g(0) - \int_0^2 g(x)d(e^x) \\ &= e^2 - \int_0^2 g(x)e^x dx. \end{aligned}$$

Example 8.3.6. *Let $f(x) = e^x$ and let g be defined on $[0,2]$ by*

$$g(x) = \begin{cases} x^2, & 0 \le x < 1 \\ \sin(x), & 1 \le x \le 2. \end{cases}$$

Show $\int f dg$ exists and evaluate it.

Solution 8.3.6. *We know that g is of bounded variation on $[1,2]$ because it is continuously differentiable with bounded derivative there. But what about on $[0,1]$? We know that the function $h(x) = x^2$ on $[0,1]$ is of bounded variation on $[0,1]$ because it is also continuously differentiable with a bounded derivative. If $\boldsymbol{\pi}$ is any partition of $[0,1]$ then we must have, using standard notation for the partition points of $\boldsymbol{\pi}$,* *that*

$$\begin{aligned}\sum_{\pi} \mid \Delta g_j \mid &= \sum_{j=0}^{p-1} \mid \Delta g_j \mid \; + \; \mid g(1) - g(x_{p-1} \mid \\ &\leq V(h,0,1) \; + \; 2 \parallel g \parallel_\infty .\end{aligned}$$

Since the choice of partition on $[0,1]$ is arbitrary, we see $g \in BV[0,1]$. Thus, combining, we have that $g \in BV[0,2]$. It then follows that $f \in RS[g,0,2]$. Now note that on $[0,1]$, we can write $g(x) = h(x)+u(x)$ where

$$u(x) = \begin{cases} 0, & 0 \leq x < 1 \\ \sin(1) - 1, & x = 1. \end{cases}$$

Then, to evaluate $\int_0^2 f dg$ we write

$$\begin{aligned}\int_0^2 f dg &= \int_0^1 f dg + \int_1^2 f dg \\ &= \int_0^1 f d(h+u) + \int_1^2 f d(\sin(x)) \\ &= \int_0^1 f d(h) + \int_0^1 f d(u) + \int_1^2 f \cos(x) dx \\ &= \int_0^1 e^x 2x dx + f(1)(u(1) - u(1^-)) + \int_1^2 e^x \cos(x) dx \\ &= \int_0^1 e^x 2x dx + e(\sin(1) - 1) + \int_1^2 e^x \cos(x) dx\end{aligned}$$

and these integrals are standard Riemann integrals that can be evaluated by parts.

8.4 Homework

Exercise 8.4.1. *Define g on $[0,2]$ by*

$$g(x) = \begin{cases} -2 & x = 0 \\ x^3 & 0 < x < 1 \\ 9/8 & x = 1 \\ x^4/4 + 1 & 1 < x < 2 \\ 7 & x = 2 \end{cases}$$

This function is from a previous exercise.

1. *Show that if $f(x) = x^4$ on $[0,2]$, then $f \in RS[g,0,2]$.*
2. *Compute $\int_0^2 f dg$.*
3. *Explain why $g \in RS[f,0,2]$.*
4. *Compute $\int_0^2 g df$.*

Exercise 8.4.2. *Define g on $[0,2]$ by*

$$g(x) = \begin{cases} -1 & x = 0 \\ x^2 & 0 < x < 1 \\ 7/4 & x = 1 \\ \sqrt{x+3} & 1 < x < 2 \\ 3 & x = 2 \end{cases}$$

This function is also from a previous exercise.

1. *Show that if $f(x) = x^2 + 5$ on $[0,2]$, then $f \in RS[g,0,2]$.*
2. *Compute $\int_0^2 f dg$.*
3. *Explain why $g \in RS[f,0,2]$.*
4. *Compute $\int_0^2 g df$.*

Exercise 8.4.3. *Let f and g be defined on $[0,4]$ by*

$$f(x) = \begin{cases} x, & x \in Q \cap [0,4] \\ 2x, & x \in Ir \cap [0,4], \end{cases} \quad g(x) = \begin{cases} 1, & 0 \le x < 1 \\ 2, & 1 \le x < 2 \\ 3, & 2 \le x < 3 \\ 4, & 3 \le x \le 4. \end{cases}$$

Does $\int f dg$ exist and if so what is its value?

Exercise 8.4.4. *Let $f(x) = x^3$ and let g be defined on $[0,3]$ by*

$$g(x) = \begin{cases} x^2, & 0 \le x \le 2 \\ x^2 + 4, & 2 < x \le 3. \end{cases}$$

Show $\int f dg$ exists and evaluate it.

Exercise 8.4.5. *Let $f(x) = x^2 + 3x + 10$ and let g be defined on $[-1,5]$ by*

$$g(x) = \begin{cases} x^3, & -1 \le x \le 2 \\ -10x^2, & 2 < x \le 5. \end{cases}$$

Show $\int f dg$ exists and evaluate it.

Exercise 8.4.6. *The following are definitions of integrands f_1, f_2 and f_3 and integrators g_1, g_2 and g_3 on $[0,2]$. For each pair of indices i,j determine if $\int_0^2 f_i dg_j$ exists. If the integral exists, compute the value and if the integral does not exist, provide a proof of its failure to exist.*

$$f_1(x) = \begin{cases} 1, & 0 \le x < 1 \\ x-1, & 1 \le x \le 2, \end{cases} \quad f_2(x) = \begin{cases} 1, & x = 0 \\ x, & 0 < x \le 2 \end{cases} \quad f_3(x) = \begin{cases} 2, & x = 0 \\ 1, & 0 < x < 1 \\ x-1, & 1 \le x \le 2 \end{cases}$$

$$g_1(x) = \begin{cases} x, & 0 \leq x < 1 \\ x+1, & 1 \leq x \leq 2, \end{cases} \quad g_2(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ x+1, & 1 < x < 2 \\ 4, & x = 2, \end{cases} \quad g_3(x) = \begin{cases} -1, & x = 0 \\ x, & 0 < x \leq 1 \\ x+1, & 1 < x < 2 \\ 4, & x = 2. \end{cases}$$

Exercise 8.4.7. *Prove*

Theorem 8.4.1. Limit Interchange Theorem For Riemann - Stieljes Integrals

Assume $g \in BV[a,b]$ and $\{f_n\} \subseteq RS[g,a,b]$ converges uniformly to f_0 on $[a,b]$. Then

(i) $f_0 \in RS[g,a,b]$,

(ii) *If $F_n(x) = \int_a^x f_n(t)dg(t)$ and $F_0(x) = \int_a^x f_0(t)dg(t)$, then F_n converges uniformly to F_0 on $[a,b]$.*

(iii)

$$\lim_n \int_a^b f_n(t)dg(t) = \int_a^b f_0(t)dg(t).$$

Exercise 8.4.8. *Let g be strictly monotone on $[a,b]$. For f_1, f_2 in $C[a,b]$, define $\omega : C[a,b] \times C[a,b] \to \Re$ by $\omega(f_1,f_2) = \int_a^b f_1(t)f_2(t)dg(t)$.*

(i) *Prove that ω is an inner product on $C[a,b]$.*

(ii) *Prove if $\omega(f,h) = 0$ for all $h \in RS[g,a,b]$, then $f = 0$.*

Chapter 9

Measurable Functions and Spaces

If you have been looking closely at how we prove the properties of Riemann and Riemann Stieljes integration, you will have noted that these proofs are intimately tied to the way we use partitions to divide the function domain into small pieces. We are now going to explore a new way to associate a given bounded function with a real number which can be interpreted as the integral.

Let X be a nonempty set. In mathematics, we study sets such as X when various properties and structures have been added. For example, we might want X to have a metric d to allow us to measure an abstract version of distance between points in X. We could study sets X which have a linear or vector space structure and if this resulting vector space possessed a norm $\| \cdot \|$, we could determine an abstract version of the magnitude of objects in X. Here, we want to look at collections of subsets of the set X and impose some conditions on the structure of these collections.

Definition 9.0.1. Sigma Algebras

Let X be a nonempty set. A family of subsets $\mathcal{S}$ is called a $\boldsymbol{\sigma}$ - algebra if

(i) $\emptyset$, $X \in \mathcal{S}$.

(ii) If $A \in \mathcal{S}$, so is A^C. We say $\mathcal{S}$ is closed under complementation or complements.

(iii) If $\{A_n\}_{n=1}^{\infty} \in \mathcal{S}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{S}$. We say $\mathcal{S}$ is closed under countable unions.

The pair $(X, \mathcal{S})$ will be called a measurable space and if $A \in \mathcal{S}$, we will call A an $\mathcal{S}$ measurable set. If the underlying $\boldsymbol{\sigma}$ - algebra is understood, we usually just say, A is a measurable subset of X.

A common tool we use in working with countable collections of sets are *De Morgan's Laws*.

Lemma 9.0.2. De Morgan's Laws

Let X be a nonempty set and $\{A_\alpha | \alpha \in \Lambda\}$ be any collection of subsets of X. Hence, the index set Λ may be finite, countably infinite or arbitrary cardinality. Then

(i)

$$\left(\cup_\alpha A_\alpha \right)^C = \cap_\alpha A_\alpha^C$$

(ii)

$$\left(\cap_\alpha A_\alpha \right)^C = \cup_\alpha A_\alpha^C$$

Proof. *This is a standard proof and is left to you as an exercise.* ■

9.1 Examples

Let's work through a series of examples of $\boldsymbol{\sigma}$ algebras.

Example 9.1.1. *Let X be any not empty set and let $\mathcal{S} = \{A | A \subseteq X\}$. This is the collection of all subsets and is sometimes called the power set of X. It is often denoted by the symbol $\mathcal{P}(X)$. This collection clearly is a $\boldsymbol{\sigma}$ algebra. Hence, $(\mathcal{P}(X), X)$ is a measurable space and all subsets of X are $\mathcal{P}(X)$ measurable.*

Example 9.1.2. *Let X be any set and $\mathcal{S} = \{\emptyset, X\}$. Then this collection is also a $\boldsymbol{\sigma}$ algebra, albeit not a very interesting one! With this $\boldsymbol{\sigma}$ algebra, X is a measurable space with only two measurable sets.*

Example 9.1.3. *Let X be the set of counting numbers and let $\mathcal{S} = \{\emptyset, \boldsymbol{O}, \boldsymbol{E}, X\}$ where $\boldsymbol{O}$ is the odd counting numbers and $\boldsymbol{E}$, the odd. It is easy to see $(\mathcal{S},\ X)$ is a measurable space.*

Example 9.1.4. *Let X be any uncountable set and let $\mathcal{S} = \{A \subseteq X | A$ is countable or A^C is countable$\}$. It is easy to see $\emptyset$ and X itself are in $\mathcal{S}$. If $A \in \mathcal{S}$, then there are two cases: A is countable and /or A^C is countable. In both cases, it is straightforward to reason that A^C is also in $\mathcal{S}$. It remains to show that $\mathcal{S}$ is closed under countable unions. To do this, assume we have a sequence of sets A_n from $\mathcal{S}$. Consider $A = \cup_n A_n$. There are several cases to consider.*

1. *If all the A_n are countable, then so is the countable union implying $A \in \mathcal{S}$.*

2. *If all the A_n are not countable, then each A_n^C is countable. Thus, $\cap_n A_n^C = (\cup_n A_n)^C$ is countable. Again, this tells us $A \in \mathcal{S}$.*

3. *If a countable number of A_n and a countable number of A_n^C are uncountable, then we have, since X is uncountable,*

$$
\begin{aligned}
\left(\cup_n A_n\right)^C &= \left(\cap_n A_n^C\right) \\
&= \left(\cap_{(A_n \text{ countable})} A_n^C\right) \cap \left(\cap_{(A_n \text{ uncountable})} A_n^C\right) \\
&= \left(\cap_{(A_n^C \text{ uncountable})} A_n^C\right) \cap \left(\cap_{(A_n^C \text{ countable})} A_n^C\right)
\end{aligned}
$$

Now, for any index n, we must have $\cap_n A_n^C \subseteq A_n^C$. Thus, since some A_n^C are countable, we must have $\cap_n A_n^C$ is countable. By De Morgan's Laws, it follows that $(\cup_n A_n)^C$ is countable. This implies $A \in \mathcal{S}$.

We conclude $(X, \mathcal{S})$ is a measurable space.

Example 9.1.5. *Let X be any nonempty set and let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two sigma - algebras of X. Let*

$$
\begin{aligned}
\mathcal{S}_3 &= \{A \subseteq X \,|\, A \in \mathcal{S}_1 \text{ and } A \in \mathcal{S}_2\} \\
&\equiv \mathcal{S}_1 \cap \mathcal{S}_2.
\end{aligned}
$$

It is straightforward to see that $(X, \mathcal{S}_3)$ is a measurable space.

Example 9.1.6. *Let X be any nonempty set. Let $\mathcal{A}$ be any nonempty collection of subsets of X. Note that $\mathcal{P}(X)$, the collection of all subsets of X, is a sigma - algebra of X and hence, $(X, \mathcal{P}(X))$ is a measurable space that contains $\mathcal{A}$. By Example 9.1.5, we know if $\mathcal{S}_1$ and $\mathcal{S}_2$ are two other sigma - algebras that contain $\mathcal{A}$, then $\mathcal{S}_1 \cap \mathcal{S}_2$ is a new sigma - algebra that also contains $\mathcal{A}$. This suggests we search for the* smallest *sigma - algebra that contains $\mathcal{A}$.*

Definition 9.1.1. The Sigma - Algebra Generated By Collection A

The sigma - algebra generated by a collection of subsets $\mathcal{A}$ in a nonempty set X, is denoted by $\sigma(\mathcal{A})$ and is defined by

$$
\sigma(\mathcal{A}) = \cap\{\mathcal{S} \,|\, \mathcal{A} \subseteq \mathcal{S}\}.
$$

Since any sigma - algebra $\mathcal{S}$ that contains $\mathcal{A}$ by definition satisfies $\sigma(\mathcal{A}) \subseteq \mathcal{S}$, it is easy to see why we interpret this generated sigma - algebra as the smallest *sigma - algebra that contains the collection $\mathcal{A}$.*

9.2 The Borel Sigma - Algebra of $\Re$

We now discuss a very important sigma algebra of subsets of the real line called the Borel sigma - algebra which is denoted by $\mathcal{B}$. Define four collections of subsets of $\Re$ as follows:

1. $\boldsymbol{A}$ is the collection of finite open intervals of the form (a, b),
2. $\boldsymbol{B}$ is the collection of finite half open intervals of the form $(a, b]$,

3. $\boldsymbol{C}$ is the collection of finite half open intervals of the form $[a, b)$ and
4. $\boldsymbol{D}$ is the collection of finite closed intervals of the form $[a, b]$.

It is possible to show that

$$\sigma(\boldsymbol{A}) = \sigma(\boldsymbol{B}) = \sigma(\boldsymbol{C}) = \sigma(\boldsymbol{A}).$$

This common sigma - algebra is what we will call the Borel sigma - algebra of $\Re$. It should be evident to you that a set can be very complicated and still be in $\mathcal{B}$. Some of these equalities will be left to you as homework exercises, but we will prove that $\sigma(\boldsymbol{A}) = \sigma(\boldsymbol{D})$. Let $\mathcal{S}$ be any sigma - algebra that contains $\boldsymbol{A}$. We know that

$$\begin{aligned} [a, b] &= (-\infty, b] \cap [a, \infty) \\ &= (b, \infty)^C \cap (-\infty, a)^C \\ &= \Big((-\infty, a) \cup (b, \infty) \Big)^C. \end{aligned}$$

In the representation of $[a, b]$ above, note we can write

$$\begin{aligned} (-\infty, a) &= \bigcup_{\lfloor a \rfloor}^{\infty} (-n, a) \\ (b, \infty) &= \bigcup_{\lceil b \rceil}^{\infty} (b, n). \end{aligned}$$

Since, $\mathcal{S}$ is a sigma - algebra containing $\boldsymbol{A}$, the unions on the right hand sides in the equations above must be in $\mathcal{S}$. This immediately tells us that $[a, b]$ is also in $\mathcal{S}$. Hence, since $[a, b]$ is arbitrary, we conclude $\boldsymbol{D}$ is contained in $\mathcal{S}$ also. Further, since is true for any sigma - algebra that contains $\boldsymbol{A}$, we have that $\boldsymbol{D} \subseteq \sigma(\boldsymbol{A})$. Thus, by definition, we can say $\sigma(\boldsymbol{D}) \subseteq \sigma(\boldsymbol{A})$.

To show the reverse containment is quite similar. Let $\mathcal{S}$ be any sigma - algebra that contains $\boldsymbol{D}$. We know that

$$\begin{aligned} (a, b) &= (-\infty, b) \cap (a, \infty) \\ &= [b, \infty)^C \cap (-\infty, a]^C \\ &= \Big((-\infty, a] \cup [b, \infty) \Big)^C. \end{aligned}$$

In the representation of (a, b) above, note we can write

$$(-\infty, a] = \bigcup_{\lfloor a \rfloor}^{\infty} [-n, a]$$

and

$$[b, \infty) = \bigcup_{\lceil b \rceil}^{\infty} [b, n].$$

Since, $\mathcal{S}$ is a sigma - algebra containing $\boldsymbol{D}$, the unions on the right hand sides in the equations above must be in $\mathcal{S}$. This immediately tells us that (a, b) is also in $\mathcal{S}$. Hence, since (a, b) is arbitrary, we conclude $\boldsymbol{A}$ is contained in $\mathcal{S}$ also. Further, since is true for any sigma - algebra that contains $\boldsymbol{A}$, we have that $\boldsymbol{A} \subseteq \sigma(\boldsymbol{D})$. Thus, by definition, we can say $\sigma(\boldsymbol{A}) \subseteq \sigma(\boldsymbol{D})$. Combining, we have the equality we seek.

9.2.1 Homework

Exercise 9.2.1. *Prove* $\sigma(\boldsymbol{A}) = \sigma(\boldsymbol{B})$.

Exercise 9.2.2. *Prove* $\sigma(\boldsymbol{B}) = \sigma(\boldsymbol{C})$.

Exercise 9.2.3. *Prove any Cantor set is in* $\mathcal{B}$.

9.3 The Extended Borel Sigma Algebra

It is often very convenient to deal with a number system that explicitly adjoins the symbols ∞ and $-\infty$ to the standard real line $\Re$. This is actually called the two - point compactification of $\Re$, but that is another story!

Definition 9.3.1. The Extended Real Number System

The extended real number systems is denoted by $\overline{\Re}$ *and is defined as the real numbers with two additional elements:*

$$\overline{\Re} \quad = \quad \Re \cup \{+\infty\} \cup \{-\infty\}.$$

We want arithmetic involving the new symbols $\pm\infty$ *to reflect our everyday experience with limits of sequences of numbers which either grow without bound positively or negatively. Hence, we use the conventions for all real numbers* x*:*

$$\begin{aligned}
(\pm\infty) + (\pm\infty) &= x + (\pm\infty) = (\pm\infty) + x = \pm\infty, \\
(\pm\infty) \cdot (\pm\infty) &= \infty, \\
(\pm\infty) \cdot (\mp\infty) &= = (\mp\infty) \cdot (\pm\infty) = -\infty, \\
x \cdot (\pm\infty) &= = (\pm\infty) \cdot x = \pm\infty \text{ if } x > 0, \\
x \cdot (\pm\infty) &= = (\pm\infty) \cdot x = 0 \text{ if } x = 0, \\
x \cdot (\pm\infty) &= = (\pm\infty) \cdot x = \mp\infty \text{ if } x < 0.
\end{aligned}$$

We can not define the arithmetic operations $(\infty) + (-\infty)$, $(-\infty) + (\infty)$ *or any the four ratios of the form* $(\pm\infty)/(\pm\infty)$.

We can now define the Borel sigma - algebra in $\overline{\Re}$. Let E be any Borel set in $\Re$. Let

$$E_1 \quad = \quad E \cup \{-\infty\}, \; E_2 = E \cup \{+\infty\}, \quad \text{and} \quad E_3 = E \cup \{+\infty\} \cup \{+\infty\}.$$

Then, we define

$$\overline{\mathcal{B}} \quad = \quad \{E, E_1, E_2, E_3 \,|\, E \in \mathcal{B}\}.$$

We leave to you the exercise of showing that $\overline{\mathcal{B}}$ is a sigma - algebra in $\overline{\Re}$.

Exercise 9.3.1. *Prove that $\overline{\mathcal{B}}$ is a sigma - algebra in $\overline{\Re}$.*

It is that open intervals in $\Re$ are in $\overline{\mathcal{B}}$, but is it true that $\overline{\mathcal{B}}$ contains arbitrary open sets? To see that it does, we must prove a characterization for the open sets of $\Re$.

Theorem 9.3.1. Open Set Characterization Lemma
If $\mathcal{U}$ is an open set in $\Re$, then there is a countable collection of disjoint open intervals $\mathcal{C} = \{(a_n, b_n)\}$ so that $\mathcal{U} = \cup_n (a_n, b_n)$.

Proof. *Since $\mathcal{U}$ is open, if $p \in \mathcal{U}$, there is an $r > 0$ so that $B(p;r) \subseteq \mathcal{U}$. Hence, $(p-r, p+r) \subseteq \mathcal{U}$ implying both $(p, p+r) \subseteq \mathcal{U}$ and $(p-r, p) \subseteq \mathcal{U}$. Let*

$$S_p \quad = \quad \{y \,|\, (p, y) \subseteq \mathcal{U}\} \quad \textit{and} \quad T_p \; = \; \{x \,|\, (x, p) \subseteq \mathcal{U}\}.$$

It is easy to see that both S_p and T_p are nonempty since $\mathcal{U}$ is open. Let $b_p = \sup S_p$ and $a_p = \inf T_p$. Clearly, b_p could be $+\infty$ and a_p could be $-\infty$.

Consider $u \in (a_p, b_p)$. From the Infimum and Supremum tolerance lemmas, we know there are points x^* and y^* so that

$$\begin{aligned} u \;<\; y^* \;\leq\; b_p \;\leq\; \infty \quad &\textit{and} \quad (p, y^*) \subseteq \mathcal{U} \\ , -\infty \;\leq\; a_p \;\leq\; x^* \;<\; u \quad &\textit{and} \quad (x^*, p) \subseteq \mathcal{U}. \end{aligned}$$

Hence, $u \in (x^*, y^*) \subseteq \mathcal{U}$ which implies $u \in \mathcal{U}$. Thus, since u in (a_p, b_p) is arbitrary, we have $(a_p, b_p) \subseteq \mathcal{U}$. if a_p or b_p were not finite, they can not be in $\Re$ and can not be in $\mathcal{U}$. However, what if either one was finite? Is it possible for the point to be in $\mathcal{U}$? We will show that in this case, the points a_p and b_p still can not lie in $\mathcal{U}$. For concreteness, let us assume that a_p is finite and in $\mathcal{U}$. Then, a_p would be an interior point of $\mathcal{U}$. Hence, there would be a radius $\rho > 0$ so that $(a_p - \rho, a_p) \subseteq \mathcal{U}$ implying $a_p - \rho \in T_p$. Thus, $\inf T_p = a_p \leq a_p - \rho$ which is not possible. Hence, $a_p \notin \mathcal{U}$. A similar argument then shows that if b_p is finite, b_p is not in $\mathcal{U}$.

Thus, we know that a_p and b_p are never in $\mathcal{U}$ and that p is always in the open interval $(a_p, b_p) \subseteq \mathcal{U}$. Let $\mathcal{F} = \{(a_p, b_p) \,|\, p \in \mathcal{U}\}$. We see immediately that

$$\mathcal{U} \quad = \quad \cup_{\mathcal{F}} \, (a_p, b_p).$$

Let (a, b) and (c, d) be any two intervals from $\mathcal{F}$ which overlap. From the definition of $\mathcal{F}$, we then know that a, b, c and d are not in $\mathcal{U}$. Then, if $a \geq d$, the two intervals would be disjoint; hence, we must have $a < d$. By the same sort of argument, it is also true that $c < b$. Hence, if c is in the intersection, we have a chain of inequalities like this:

$$a \;<\; c \;<\; q \;<\; b \;<\; d.$$

Next, since $a \notin \mathcal{U}$, we see $a \leq c$ since $(c, d) \subseteq \mathcal{U}$. Further, since $c \notin \mathcal{U}$ and $(a, b) \subseteq \mathcal{U}$, it follows that $c \leq a$. Combining, we have $a = c$. A similar argument shows that $b = d$. Hence, $(a, b) \cap (c, d) \neq \emptyset$ implies that $(a, b) = (c, d)$. Thus, two interval I_p and I_q in $\mathcal{F}$ are either the same or disjoint. We conclude

$$\mathcal{U} = \bigcup_{(disjoint\ I_p \in \mathcal{F})} I_p.$$

Let $\mathcal{F}_0$ be this collection of disjoint intervals from $\mathcal{F}$. Each I_p in $\mathcal{F}_0$ contains a rational number r_p. By definition, it then follows that if I_p and I_q are in $\mathcal{F}_0$, then $r_p \neq r_q$. The set of these rational numbers is countable and so we can label them using an enumeration r_n. Label the interval I_p which contains r_n as I_n. Then, we have

$$\mathcal{U} = \bigcup_{n=1}^{\infty} I_n,$$

which is the desired result. ■

9.4 Measurable Functions

Let $f : \Re \rightarrow \Re$ be a continuous function. let $\mathcal{O}$ be an open subset of $\Re$. By Theorem 9.3.1, we know that we can write

$$\mathcal{O} = \bigcup_n (a_n, b_n)$$

where the (a_n, b_n) are mutually disjoint finite open intervals of $\Re$. It follows immediately that $\mathcal{O}$ is in the Borel sigma - algebra $\mathcal{B}$. Now consider the inverse image of $\mathcal{O}$ under f, $f^{-1}(\mathcal{O})$. If $p \in f^{-1}(\mathcal{O})$, then $f(p) \in \mathcal{O}$. Since $\mathcal{O}$ is open, $f(p)$ must be an interior point. Hence, there is a radius $r > 0$ so that $(f(p) - r, f(p) + r) \subseteq \mathcal{O}$. Since f is continuous at p, there then is a $\delta > 0$ so that $f(x) \in ((f(p) - r, f(p) + r)$ if $x \in (p - \delta, p + \delta)$. This tells us that $(p - \delta, p + \delta) \subseteq f^{-1}(\mathcal{O})$. Since p was arbitrarily chosen, we conclude that $f^{-1}(\mathcal{O})$ is an open set.

We see that if f is continuous on $\Re$, then $f^{-1}(\mathcal{O})$ is in the Borel sigma - algebra for any open set $\mathcal{O}$ in $\Re$. We can then say that $f^{-1}(\alpha, \infty)$ is in $\mathcal{B}$ for all $\alpha > 0$. This suggests that an interesting way to generalize the notion of *continuity* might be to look for functions f on an arbitrary nonempty set X with sigma - algebra $\mathcal{S}$ satisfying $f^{-1}(\mathcal{O}) \in \mathcal{S}$ for all open sets $\mathcal{O}$. Further, by our last remark, it should be enough to ask that $f^{-1}((\alpha, \infty)) \in \mathcal{S}$ for all $\alpha \in \Re$. This is exactly what we will do. It should be no surprise to you that functions f satisfying this new definition will not have to be continuous!

Definition 9.4.1. The Measurability of a Function

Let X be a nonempty set and $\mathcal{S}$ be a sigma - algebra of subsets of X. We say that $f : X \rightarrow \Re$ is a $\mathcal{S}$ - measurable function on X or simply $\mathcal{S}$ measurable if

$$\forall\, \alpha \ > \ 0, \ \{x \in X \,|\, f(x) \ >: \alpha\} \ \in \ \mathcal{S}.$$

We can easily prove that there are equivalent ways of proving a function is measurable.

Lemma 9.4.1. Equivalent Conditions For The Measurability of a Function
Let X be a nonempty set and $\mathcal{S}$ be a sigma - algebra of subsets of X. The following statements are equivalent:

(i): $\forall \alpha > 0,\ A_\alpha = \{x \in X \mid f(x) > \alpha\} \in \mathcal{S}$,

(ii): $\forall \alpha > 0,\ B_\alpha = \{x \in X \mid f(x) \leq \alpha\} \in \mathcal{S}$,

(iii): $\forall \alpha > 0,\ C_\alpha = \{x \in X \mid f(x) \geq \alpha\} \in \mathcal{S}$,

(iv): $\forall \alpha > 0,\ D_\alpha = \{x \in X \mid f(x) < \alpha\} \in \mathcal{S}$.

Proof.
(i) $\Rightarrow$ (ii):

Subproof. *If $A_\alpha \in \mathcal{S}$, then its complement is in $\mathcal{S}$ also. Since $B_\alpha = A_\alpha^C$, (ii) follows.* □

(ii) $\Rightarrow$ (i):

Subproof. *If $B_\alpha \in \mathcal{S}$, then its complement is in $\mathcal{S}$ also. Since $A_\alpha = B_\alpha^C$, (i) follows.* □

(iii) $\Leftrightarrow$ (iv):

Subproof. *Since $C_\alpha = D_\alpha^C$ and $D_\alpha = C_\alpha^C$, arguments similar to those of the previous cases can be applied.* □

Hence, if we show (i) $\Leftrightarrow$ (iii), we will be done. (i) $\Rightarrow$ (iii):

Subproof. *By (i), $A_{\alpha-1/n} \in \mathcal{S}$ for all n. We know*

$$C_\alpha = \bigcap_n A_{\alpha-1/n} = \bigcap_n \{x | f(x) > \alpha - 1/n\}$$

We also know $A^C_{\alpha-1/n}$ is measurable and so $\cup_n A^C_{\alpha-1/n}$ is also measurable. Thus, the complement of $\cup_n A^C_{\alpha-1/n}$ is also measurable. Then, by De Morgan's Laws, $C_\alpha = \cap_n A_{\alpha-1/n}$ is measurable. □

(iii) $\Rightarrow$ (i):

Subproof. *Note, $C_{\alpha+1/n} \in \mathcal{S}$ for all n and so*

$$A_\alpha = \bigcup_n C_{\alpha+1/n} = \bigcup_n \{x | f(x) \geq \alpha + 1/n\}$$

is also measurable. □

We conclude all four statements are equivalent. ■

9.4.1 Examples

Example 9.4.1. *Any constant function f on a nonempty set X with given sigma - algebra $\mathcal{S}$ is measurable as if $f(x) = c$ for some $c \in \Re$, then*

$$\{x | f(x) > \alpha\} = \begin{cases} \emptyset \in \mathcal{S} & \alpha \geq c \\ X \in \mathcal{S} & \alpha < c \end{cases}$$

Example 9.4.2. *Let X be a nonempty set X with given sigma - algebra $\mathcal{S}$. Let $E \in \mathcal{S}$ be given. Define*

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

Then I_E is measurable. Note

$$\{x | I_E(x) > \alpha\} = \begin{cases} \emptyset \in \mathcal{S} & \alpha \geq 1 \\ E \in \mathcal{S} & 0 \leq \alpha < 1 \\ X \in \mathcal{S} & \alpha < 0 \end{cases}$$

Example 9.4.3. *Let $X = \Re$ and $\mathcal{S} = \mathcal{B}$. Then, if $f : \Re \rightarrow \Re$ is continuous, f is measurable by the arguments we made at the beginning of this section. More generally, let $f : [a, b] \rightarrow \Re$ be continuous on $[a, b]$. Then, extend f to $\Re$ as $\hat{f}$ defined by*

$$\hat{f} = \begin{cases} f(a) & x < a \\ f(x) & a \leq x \leq b \\ f(b) & x > b \end{cases}$$

Then $\hat{f}$ is continuous on $\Re$ and measurable with $\hat{f}^{-1}(\alpha, \infty) \in \mathcal{B}$ for all α. It is not hard to show that

$$\mathcal{B} \cap [a, b] = \{E \subseteq [a, b] \mid E \in \mathcal{B}\}$$

is a sigma -algebra of the set $[a, b]$. Further, the standard arguments for f continuous on $[a, b]$ show us that $f^{-1}(\alpha, \infty) \in \mathcal{B} \cap [a, b]$ for all α. Hence, a continuous f on the interval $[a, b]$ will be measurable with respect to the sigma - algebra $\mathcal{B} \cap [a, b]$.

We can argue is a similar fashion for functions continuous on intervals of the form $(a, b]$, $[a, b)$ and (a, b) whether a and b is finite or not.

Example 9.4.4. *If $X = \Re$ and $\mathcal{S} = \mathcal{B}$, then any monotone function if Borel measurable. To see this, note we can restrict our attention to monotone increasing functions as the argument is quite similar for monotone decreasing.*

$$\{x | f(x) > \alpha\} = \begin{cases} \emptyset \in \mathcal{S} & f(x) \leq \alpha \,\forall x \\ X \in \mathcal{S} & f(x) > \alpha \,\forall x. \end{cases}$$

Hence, it is enough to consider the cases where f takes on the value α without a jump at the point x_0 or f has a jump across the value α at x_0. In the first case, since f is monotone increasing and $f(x_0) = \alpha$, $f^{-1}(\alpha, \infty) = (x_0, \infty) \in \mathcal{B}$. On the other hand, if f has a jump at x_0 across the value α, , then $f(x_0^-) \neq f(x_0^+)$ and $\alpha \in [f(x_0^-), f(x_0^+)]$. there are three possibilities:

- *(i): $f(x_0^-) = f(x_0) < f(x_0^+)$: If $\alpha = f(x_0)$, then since f is monotone, $f^{-1}(\alpha, \infty) = (x_0, \infty)$. If $f(x_0) < \alpha < f(x_0^+)$, we again have $f^{-1}(\alpha, \infty) = (x_0, \infty)$. Finally, if $\alpha = f(x_0^+)$, we have $f^{-1}(\alpha, \infty) = [x_0, \infty)$. In all cases, these inverse images are in $\mathcal{B}$.*
- *(ii): $f(x_0^-) < f(x_0) < f(x_0^+)$: A similar analysis shows that all the possible inverse images are Borel sets.*
- *(iii): $f(x_0^-) < f(x_0) = f(x_0^+)$: we handle the arguments is a similar way.*

We conclude that in all cases, $f^{-1}(\alpha, \infty) \in \mathcal{B}$ and hence f is measurable.

Note, the analysis of the previous example could be employed here also to show that a monotone function defined on an interval such as $[a, b]$, (a, b) and so forth is Borel measurable with respect to the restricted sigma - algebra $\mathcal{B} \cap [a, b]$ etc.

Exercise 9.4.1. *Let f be piecewise continuous on $[a, b]$. Prove that f is measurable with respect to the restricted Borel sigma - algebra $\mathcal{B} \cap [a, b]$. Recall, a function is piecewise continuous on $[a, b]$ if there are a finite number of points x_i in $[a, b]$ where f is not continuous.*

Comment 9.4.1. *For convenience, we will start using a more abbreviated notation for sets like $\{x \in X \mid f(x) > \alpha\}$; we will shorten this to $\{f(x) > \alpha\}$ or $(f(x) > \alpha)$ in our future discussions.*

9.5 Properties of Measurable Functions

We now want to see how we can build new measurable functions from old ones we know.

Lemma 9.5.1. Properties of Measurable Functions

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Then if f and g are $\mathcal{S}$ measurable, so are

- *(i): cf for all $c \in \Re$.*
- *(ii): f^2.*
- *(iii): $f + g$.*
- *(iv): fg.*
- *(v): $\mid f \mid$.*

Proof.
(i):

Subproof. *If $c = 0$, $cf = 0$ and the result is clear. If $c > 0$, then $(cf(x) > \alpha) = (f(x) > \alpha/c)$ which is measurable as f is measurable. If $c < 0$, a similar argument holds.* □

(ii):

Subproof. *If $\alpha < 0$, then $(f^2(x) > \alpha) = X$ which is in $\mathcal{S}$. Otherwise, if $\alpha \geq 0$, then*

$$(f^2(x) > \alpha) \ = \ (f(x) > \sqrt{\alpha}) \cup (f(x) < -\sqrt{\alpha}),$$

and both of these sets are measurable since f is measurable. The conclusion follows. □

(iii):

Subproof. *If $r \in Q$, let $S_r = (f(x) > r) \cap (g(x) > \alpha - r)$ which is measurable since f and g are measurable. We claim that*

$$(f(x) + g(x) > \alpha) \ = \ \bigcup_{r \in Q} S_r.$$

To see this, let x satisfy $f(x) + g(x) > \alpha$. Thus, $f(x) > \alpha - g(x)$. Since the rationals are dense in $\Re$, we see there is a rational number r so that $f(x) > r > g(x) - \alpha$. This clearly implies that $f(x) > \alpha$ and $g(x) > \alpha - r$ and so $x \in S_r$. Since our choice of x was arbitrary, we have shown that

$$(f(x) + g(x) > \alpha) \ \subseteq \ \bigcup_{r \in Q} S_r.$$

The converse is easier as if $x \in S_r$, it follows immediately that $f(x) + g(x) > \alpha$.

Since S_r is measurable for each r and the rationals are countable, we see $(f(x) + g(x) > \alpha)$ is measurable. □

(iv):

Subproof. *To prove this result, note that $fg = (1/4)\Big((f+g)^2 - (f-g)^2\Big)$ and all the individual pieces are measurable by (iii) and (i).* □

(v):

Subproof. *If $\alpha < 0$, $(f(x) > \alpha) = X$ which is measurable. On the other hand, if $\alpha \geq 0$,*

$$(f^2(x) > \alpha) \ = \ (f(x) > \alpha) \cup (f(x) < -\alpha),$$

which implies the measurability of $|\ f\ |$. □

■

We can also prove another characterization of the measurability of f.

Lemma 9.5.2. A Function is Measurable If and Only If Its Positive and Negative Parts Are Measurable

Let X be a nonempty set and $\mathcal{S}$ be a sigma - algebra on X. Then $f : X \rightarrow \Re$ is measurable if and only if f^+ and f^- are measurable, where

$$f^+(x) \ = \ \max\{f(x), 0\}, \quad \textit{and} \quad f^+(x) \ = \ -\min\{f(x), 0\}.$$

Proof. *We note* f $f^+ - f^-$ *and* $|\ f\ | = f^+ + f^-$. *Thus,* $f^+ = (1/2)\Big(|\ f\ | + f\Big)$ *and* $f^- = (1/2)\Big(|\ f\ | - f\Big)$. *Hence, if* f *is measurable, by Lemma 9.5.1 (i), (iii) and (v),* f^+ *and* f^- *are also measurable. Conversely, if* f^+ *and* f^- *are measurable,* $f = f^+ - f^-$ *is measurable as well.* ■

9.6 Extended Valued Measurable Functions

We now extend these ideas to functions which are extended real valued.

Definition 9.6.1. The Measurability Of An Extended Real Valued Function

Let X *be a nonempty set and* $\mathcal{S}$ *be a sigma - algebra on* X. *Let* $f : X \to \overline{\Re}$. *We say* f *is* $\mathcal{S}$ *measurable if* $(f(x) > \alpha)$ *is in* $\mathcal{S}$ *for all* α *in* $\Re$.

Comment 9.6.1. *If the extended valued function* f *is measurable, then* $(f(x) = +\infty) = \cap_n (f(x) > n)$ *is measurable. Also, since* $(f(x) = -\infty) = \Big(\cup_n (f(x) > -n)\Big)^C$, *it is measurable also.*

We can then prove an equivalence theorem just like before.

Lemma 9.6.1. Equivalent Conditions For The Measurability of an Extended Real Valued Function

Let X *be a nonempty set and* $\mathcal{S}$ *be a sigma - algebra of subsets of* X. *The following statements are equivalent:*

(i): $\forall \alpha > 0,\ A_\alpha = \{x \in X \mid f(x) > \alpha\} \in \mathcal{S}$,

(ii): $\forall \alpha > 0,\ B_\alpha = \{x \in X \mid f(x) \leq \alpha\} \in \mathcal{S}$,

(iii): $\forall \alpha > 0,\ C_\alpha = \{x \in X \mid f(x) \geq \alpha\} \in \mathcal{S}$,

(iv): $\forall \alpha > 0,\ D_\alpha = \{x \in X \mid f(x) < \alpha\} \in \mathcal{S}$.

Proof. *The proof follows that of Lemma 9.4.1* ■

The collection of all extended valued measurable functions is important to future work. We make the following definition:

Definition 9.6.2. The Set of Extended Real Valued Measurable Functions

Let X *be a nonempty set and* $\mathcal{S}$ *be a sigma - algebra of subsets of* X. *We denote by* $M(X, \mathcal{S})$ *the set of all extended real valued measurable functions on* X. *Thus,*

$$M(X, \mathcal{S}) \quad = \quad \{f : X \to \overline{\Re} \mid f \ \text{ is } \ \mathcal{S} \text{ measurable}\}.$$

It is also easy to prove the following equivalent definition of measurability for extended valued functions.

Lemma 9.6.2. Extended Valued Measurability In Terms Of The Finite Part Of The Function

Let X *be a nonempty set and* $\mathcal{S}$ *be a sigma - algebra of subsets of* X. *Then* $f \in M(X, \mathcal{S})$ *if and only if (i):* $(f(x) = +\infty) \in \mathcal{S}$, *(ii):* $(f(x) = -\infty) \in \mathcal{S}$ *and (iii):* f_1 *is measurable where*

$$f_1(x) \quad = \quad \begin{cases} f(x) & x \notin (f(x) = +\infty) \ \cup \ (f(x) = -\infty), \\ 0 & x \in (f(x) = +\infty) \ \cup \ (f(x) = -\infty). \end{cases}$$

Proof. *By Comment 9.6.1, if f is measurable, (i) and (ii) are true. Now, if $\alpha \geq 0$ is given, we see*

$$(f_1(x) > \alpha) \quad = \quad (f(x) > \alpha) \quad (f(x) = +\infty) \; = \; (f(x) > \alpha) \cap (f(x) = +\infty)^C,$$

which is a measurable set. On the other hand, if $\alpha < 0$, then

$$(f_1(x) > \alpha) \quad = \quad (f(x) > \alpha) \cup (f(x) = -\infty),$$

which is measurable as well. We conclude f_1 is measurable. Conversely, if (i), (ii) and (iii) hold, then if $\alpha \geq 0$, we have

$$(f(x) > \alpha) \quad = \quad (f_1(x) > \alpha) \cap (f(x) = +\infty),$$

and if $\alpha < 0$,

$$(f(x) > \alpha) \quad = \quad (f_1(x) > \alpha) \quad (f(x) = -\infty),$$

implying both sets are measurable. Thus, f is measurable. ∎

Example 9.6.1. *Let X be a nonempty set X with given sigma - algebra $\mathcal{S}$. Let $E \in \mathcal{S}$ be given. Define the extended value characteristic function*

$$I_E(x) \quad = \quad \begin{cases} \infty & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Then I_E is measurable. Note

$$\{x | I_E(x) > \alpha\} \quad = \quad \begin{cases} E \in \mathcal{S} & \alpha \geq 0 \\ X \in \mathcal{S} & \alpha < 0 \end{cases}$$

Note also that if we define

$$I_E(x) \quad = \quad \begin{cases} \infty & \text{if } x \in E \\ -\infty & \text{if } x \notin E \end{cases}$$

Then I_E is measurable. We have

$$\{x | I_E(x) > \alpha\} \quad = \quad \begin{cases} E \in \mathcal{S} & \alpha \geq 0 \\ E \in \mathcal{S} & \alpha < 0 \end{cases}$$

Finally, $(I_E(x) = +\infty) = E$ and $(I_E(x) = -\infty) = E^C$ are both measurable and the f_1 type function used in Lemma 9.6.2 here is $(I_E)_1(x) = 0$ always.

It is straightforward to prove these properties:

Lemma 9.6.3. Properties of Extended Valued Measurable Functions
Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Then if f and g are in $M(X, \mathcal{S})$, so are

(i): cf for all $c \in \Re$.

(ii): f^2.

(iii): $f + g$, as long as we restrict the domain of $f + g$ to be E_{fg} where

$$E_{fg}^C = \Big((f(x) = +\infty) \cap (g(x) = -\infty) \cup (f(x) = -\infty) \cap (g(x) = +\infty) \Big)^C.$$

We usually define $(f + g)(x) = 0$ on E_{fg}. Note E_{fg} is measurable since f and g are measurable functions.

(iv): $\mid f \mid$, f^+ and f^-.

Proof. *These proofs are similar to those shown in the proof of Lemma 9.5.1. However, let's look at the details of the proof of (ii). We see that our definition of addition of the extended real valued sum means that*

$$(f + g)(x) = \Big(f + g \Big) I_{E_{fg}^C}.$$

Define h by

$$h(x) = \Big(f + g \Big) I_{E_{fg}^C}(x).$$

Let α be a real number. Then

$$(h(x) > \alpha) = \begin{cases} (f(x) + g(x) > \alpha) \bigcap E_{fg}^C & \alpha \geq 0 \\ \Big((f(x) + g(x) > \alpha) \bigcap E_{fg}^C \Big) \cup E_{fg} & \alpha < 0 \end{cases}$$

Similar to what we did in Lemma 9.5.1, for $r \in Q$, let

$$S_r = (f(x) > r) \cap (g(x) > \alpha - r) \cap E_{fg}^C$$

which is measurable since f and g are measurable. We claim that

$$(f(x) + g(x) > \alpha) \cap E_{fg}^C = \bigcup_{r \in Q} S_r.$$

To see this, let x satisfy $f(x) + g(x) > \alpha$. Thus, $f(x) > \alpha - g(x)$. Since the rationals are dense in $\Re$, we see there is a rational number r so that $f(x) > r > g(x) - \alpha$. This clearly implies that $f(x) > \alpha$ and $g(x) > \alpha - r$ and so $x \in S_r$. Since our choice of x was arbitrary, we have shown that

$$(f(x) + g(x) > \alpha) \subseteq \bigcup_{r \in Q} S_r.$$

The converse is easier as if $x \in S_r$, it follows immediately that $f(x)+g(x)$ is defined and $f(x)+g(x) > \alpha$.

Since S_r is measurable for each r and the rationals are countable, we see $(f(x)+g(x) > \alpha) \cap E_{fg}^C$ is measurable. ∎

To prove that products of extended valued measurable functions are also measurable, we have to use a pointwise limit approach.

Lemma 9.6.4. Pointwise Infimums, Supremums, Limit Inferiors and Limit Superiors are Measurable

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $(f_n) \subseteq M(X, \mathcal{S})$. Then

(i): If $f(x) = \inf_n f_n(x)$, then $f \in M(X, \mathcal{S})$.

(ii): If $F(x) = \sup_n f_n(x)$, then $F \in M(X, \mathcal{S})$.

(iii): If $f^*(x) = \liminf_n f_n(x)$, then $f^* \in M(X, \mathcal{S})$.

(iv): If $F^*(x) = \limsup_n f_n(x)$, then $F^* \in M(X, \mathcal{S})$.

Proof. *It is straightforward to see that $(f(x) \geq \alpha) \cap_n (f_n(x) \geq \alpha)$ and $(F(x) \geq \alpha) \cup_n (f_n(x) > \alpha)$ and hence, are measurable for all α. It follows that f and F are in $M(X, \mathcal{S})$ and so (i) and (ii) hold. Next, recall from classical analysis that at each point x,*

$$\begin{aligned} \liminf(f_n(x)) &= \sup_n \inf_{k \geq n} f_k(x), \\ \limsup(f_n(x)) &= \inf_n \sup_{k \geq n} f_k(x). \end{aligned}$$

Now let $z_n(x) = \inf_{k \geq n} f_k(x)$ and $w_n(x) = \sup_{k \geq n} f_k(x)$. Applying (i) to z_n, we have $z_n \in M(X, \mathcal{S})$ and applying (ii) to w_n, we have $w_n \in M(X, \mathcal{S})$. Then apply (i) and (ii) to $\sup_n z_n$ and $\inf w_n$, respectively, to get the desired result. ∎

This leads to an important result.

Theorem 9.6.5. Pointwise Limits of Measurable Functions Are Measurable

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $(f_n) \subseteq M(X, \mathcal{S})$ and let $f : X \to \overline{\Re}$ be a function such that $f_n \to f$ pointwise on X. Then $f \in M(X, \mathcal{S})$.

Proof. *We know that $\liminf_n f_n(x) = \limsup_n f_n(x) = \lim_n f_n(x)$. Thus, by Lemma 9.6.4, we know that f is measurable.* ∎

Comment 9.6.2. *This is a huge result. We know from classical analysis that the pointwise limit of continuous functions need not be continuous (e.g. let $f_n(t) = t^n$ on $[0, 1]$). Thus, the closure of a class of functions which satisfy a certain property (like continuity) under a limit operation is not always guaranteed. We see that although measurable functions are certainly not as* smooth *as we would like, they are well behaved enough to be closed under pointwise limits!*

We now show that $M(X, \mathcal{S})$ is closed under multiplication.

Lemma 9.6.6. Products of Measurable Functions Are Measurable
Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $f, g \in M(X, \mathcal{S})$. Then $fg \in M(X, \mathcal{S})$.

Proof. *Let f_n, the* truncation *of f, be defined by*

$$f_n(x) = \begin{cases} f(x) & \mid f(x) \mid \leq n \\ n & f(x) > n \\ -n & f(x) < -n \end{cases}$$

We define the truncation of g, g_n, is a similar way. We can easily show f_n and g_m are measurable for any n and m. We only show the argument for f_n as the argument for g_m is identical. Let α be a given real number. Then

$$(f_n(x) > \alpha) = \begin{cases} \emptyset & \alpha \geq n, \\ (f(x) > n) \cup (\alpha < f(x) \leq n) & 0 \leq \alpha < n, \\ (f(x) > n) \cup (\alpha < f(x) \leq n) & -n < \alpha < 0, \\ X & \alpha \leq -n. \end{cases}$$

It is easy to see all of these sets are in $\mathcal{S}$ since f is measurable. Thus, each real valued f_n is measurable.

It then follows by Lemma 9.5.1 that $f_n g_m$ is also measurable. Note we are using the definition of measurability for real valued functions here. Next, an easy argument shows that at each x,

$$f(x) = \lim_n f_n(x) \quad \textit{and} \quad g(x) = \lim_m g_m(x)$$

It then follows that

$$f(x)\, g_m(x) = \lim_n \Big(f_n(x) g_m(x) \Big)$$

Using Theorem 9.6.5, we see $f g_m$ is measurable. Then, noting

$$f(x)\, g(x) = \lim_m \Big(f(x) g_m(x) \Big)$$

another application of Theorem 9.6.5 establishes the result. ∎

Lemma 9.6.7. Continuous Functions Of Finite Measurable Functions Are Measurable
Let X be nonempty and $(X, \mathcal{S}, \mu)$ be a measure space. Let $f \in M(X, \mathcal{S})$ be finite. Let $\phi : \Re \rightarrow \Re$ be continuous. Then $\phi \circ f$ is measurable.

Proof. *Let α be in $\Re$. We claim*

$$\Big(\phi \circ f \Big)^{-1} (\alpha, \infty) = f^{-1} \Big(\phi^{-1}(\alpha, \infty) \Big).$$

First, let x be in the right hand side. Then,

$$\begin{aligned} f(x) \in \phi^{-1}(\alpha, \infty) &\Rightarrow \phi\Big(f(x)\Big) \in (\alpha, \infty) \\ &\Rightarrow x \in \Big(\phi \circ f\Big)^{-1}(\alpha, \infty). \end{aligned}$$

Conversely, if x is in the left hand side, then

$$\begin{aligned} \Big(\phi \circ f\Big)(x) \in (\alpha, \infty) &\Rightarrow f(x) \in \phi^{-1}(\alpha, \infty) \\ &\Rightarrow x \in f^{-1}\Big(\phi^{-1}(\alpha, \infty)\Big). \end{aligned}$$

Since ϕ is continuous, $G = \phi^{-1}(\alpha, \infty)$ is an open set. Finally, since f is measurable, $f^{-1}(G)$ is in $\mathcal{S}$. We conclude that $\phi \circ f$ is measurable, since our choice of α is arbitrary. ■

Our final results in this section are a standard approximation result and a consequence.

Theorem 9.6.8. The Approximation Of Non negative Measurable Functions By Monotone Sequences

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $f \in M(X, \mathcal{S})$ which is non negative. Then there is a sequence $(\phi_n) \subseteq M(X, \mathcal{S})$ so that

(i): $0 \leq \phi_n(x) \leq \phi_{n+1}$ for all x and for all $n \geq 1$.

(ii): $\phi_n(x) \leq f(x)$ for all x and n and $f(x) = \lim_n \phi_n(x)$.

(iii): Each ϕ_n has a finite range of values.

Proof. *Pick a positive integer n. Let*

$$E_{k,n} = \begin{cases} \{x \in X \mid \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\}, & \text{for } 0 \leq k \leq n2^n - 1 \\ \{x \in X \mid n \leq f(x)\}, & \text{for } k = n2^n \end{cases}$$

You should draw some of these sets for a number of choices of non negative functions f to get a feel for what they mean. Once you have done this, you will see that this definition slices the $[0, n]$ range of f into $n2^n$ slices each of height 2^{-n}. The last set, $E_{n2^n,n}$ is the set of all points where $f(x)$ exceeds n. This gives us a total of $n2^n + 1$ sets. It is clear that $X = \cup_k E_{k,n}$ and that each of these sets are disjoint from the others. Now define the functions ϕ_n by

$$\phi_n(x) = \frac{k}{2^n}, \; x \in E_{k,n}.$$

It is evident that ϕ_n only takes on a finite number of values and so (iii) is established. Also, since f is measurable, we know each $E_{k,n}$ is measurable. Then, given any real numberα, the set $(\phi_n(x) > \alpha)$ is either empty or consists of a union of the finite number of sets $E_{k,n}$ with the property that $\alpha > (k/2^n)$. Thus, $(\phi_n(x) > \alpha)$ is measurable for all α. We conclude each ϕ_n is measurable. If $f(x) = +\infty$, then by definition, $\phi_n(x) = n$ for all n and we have $f(x) = \lim_n \phi_n(x)$. Note, the ϕ_n values are strictly monotonically increasing which shows (i) and (ii) both hold in this case.

On the other hand, if $f(x)$ is finite, let n_0 be the first integer with $n_0 - 1 \leq f(x) < n_0$. Then, we must have $\phi_1(x) = 1$, $\phi_2(x) = 2$ and so forth until we have $\phi_{n_0-1} = n_0 - 1$. These first values are monotone increasing. We also know from the definition of ϕ_{n_0} that there is a k_0 so that

$$\frac{k_0}{2^{n_0}} \leq f(x) < \frac{k_0+1}{2^{n_0}}.$$

Thus, $0 \leq f(x) - \phi_{n_0}(x) < 2^{-n_0}$. Now consider the function ϕ_{n_0+1}. We know

$$\begin{aligned} f(x) &\in [\frac{k_0}{2^{n_0}}, \frac{k_0+1}{2^{n_0}}) \\ &= [\frac{2k_0}{2^{n_0+1}}, \frac{2k_0+1}{2^{n_0+1}}) \cup [\frac{2k_0+1}{2^{n_0+1}}, \frac{2k_0+2}{2^{n_0+1}}) \end{aligned}$$

If $f(x)$ lands in the first interval above, we have

$$\phi_{n_0+1}(x) = \frac{2k_0}{2^{n_0+1}} = \frac{k_0}{2^{n_0}} = \phi_{n_0}(x)$$

and if $f(x)$ is in the second interval, we have

$$\phi_{n_0+1}(x) = \frac{2k_0+1}{2^{n_0+1}} > \frac{k_0}{2^{n_0}} = \phi_{n_0}(x).$$

In both cases, we have $\phi_{n_0}(x) \leq \phi_{n_0+1}(x)$. We also have immediately that $0 \leq f(x) - \phi_{n_0+1}(x) < 2^{-n_0-1}$.

The argument for $n_0 + 2$ and so on in quite similar and is omitted. This establishes (i) for this case. In general, we have $0 \leq f(x) - \phi_k(x) < 2^{-k}$ for all $k \geq n_0$. This implies that $f(x) = \lim_n \phi_n(x)$ which establishes (ii). ■

Lemma 9.6.9. Continuous Functions Of Measurable Functions Are Measurable
Let X be nonempty and $(X, \mathcal{S}, \mu)$ be a measure space. Let $f \in M(X, \mathcal{S})$. Let $\phi : \Re \to \Re$ be continuous and assume that $\lim_n \phi(n)$ and $\lim_n \phi(-n)$ are well defined extended value numbers. Then $\phi \circ f$ is measurable.

Proof. *Assume first that f is non negative. Then by Theorem 9.6.8, there is a sequence of finite non negative increasing functions (f_n) which are measurable and satisfy $f_n \uparrow f$. Let E be the set of points where f is finite. Then,*

$$\lim_n f_n(x) = \begin{cases} f(x) & x \in E^C \\ \lim_n n & x \in E. \end{cases}$$

Thus, since ϕ is continuous,

$$\lim_n \phi\Big(f_n(x)\Big) = \begin{cases} \phi(f(x)) & x \in E^C \\ \phi(\lim_n n) & x \in E. \end{cases}$$

We have assumed that $\lim_n \phi(n)$ is a well defined number β in $[\infty, \infty]$. Thus, if β is finite, we have

$$\lim_n \phi\Big(f_n(x)\Big) = \phi\Big(f\, I_{E^C}\Big) + \beta\, I_E$$

which is measurable since the first part is measurable by Lemma 9.6.7 and the second part is measurable since E is a measurable set by Lemma 9.6.2. If $\beta = \infty$, we have

$$\lim_n \phi\Big(f_n(x)\Big) = \begin{cases} \phi(f(x)) & x \in E^C \\ \infty & x \in E. \end{cases}$$

Now apply Lemma 9.6.2. Since E is measurable and f_1 defined by

$$f_1(x) = \begin{cases} \phi(f(x)) & x \in E^C \\ 0 & x \in E, \end{cases}$$

is measurable, we see $\lim_n \phi\Big(f_n(x)\Big)$ is measurable. A similar argument holds if $\beta = -\infty$. We conclude that if f is non negative, $\phi \circ f$ interpreted as above is a measurable function.

Thus, if f is arbitrary, the argument above shows that $\phi \circ f^+$ and $\phi \circ f^-$ are measurable. This implies that $\phi \circ f = \phi \circ (f^+ - f^-)$ is measurable when interpreted right. ■

9.7 Homework

Exercise 9.7.1. *If a, b and c are real numbers, define the value in the middle, $mid(a,b,c)$ by*

$$mid(a,b,c) = \inf\{\,\sup\{a,b\},\ \sup\{a,c\},\ \sup\{b,c\}\,\}.$$

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $f_1, f_2, f_3 \in M(X, \mathcal{S})$. Prove the function h defined pointwise by $h(x) = mid(f_1(x), f_2(x), f_3(x))$ is measurable.

Exercise 9.7.2. *Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $f \in M(X, \mathcal{S})$ and $A > 0$. Define f_A by*

$$f_A(x) = \begin{cases} f(x), & \mid f(x) \mid \leq A \\ A, & f(x) > A \\ -A, & f(x) < -A \end{cases}$$

Prove f_A is measurable.

Exercise 9.7.3. *Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra on X. Let $f \in M(X, \mathcal{S})$ and assume there is a positive K so that $0 \leq f(x) \leq K$ for all x. Prove the sequence ϕ_n of functions given in Theorem 9.6.8 converges uniformly to f on X.*

Exercise 9.7.4. *Let X and Y be nonempty sets and let $f : X \to Y$ be given. Prove that if $\mathcal{T}$ is a sigma - algebra of subsets of Y, then $\{f^{-1}(E) \mid E \in \mathcal{T}\}$ is a sigma - algebra of subsets of X.*

Exercise 9.7.5. *Let $(X, \mathcal{S})$ be a measurable space. Let (μ_n) be a sequence of measures on $\mathcal{S}$ with $\mu_n(X) \leq 1$ for all n. Define λ on $\mathcal{S}$ by*

$$\lambda(E) = \sum_{n=1}^{\infty} 1/2^n\, \mu_n(E)$$

for all measurable E. Prove λ is a measure on $\mathcal{S}$.

Chapter 10

Measure And Integration

Once we have a nonempty set X with a given sigma - algebra $\mathcal{S}$, we can develop an abstract version of integration. To motivate this, consider the Borel sigma - algebra on $\overline{\Re}$, $\overline{\mathcal{B}}$. We know how to develop and use an integration theory that is based on finite intervals of the form $[a, b]$ for bounded functions. Hence, we have learned to understand and perform integrations of the form $\int_a^b f(t)dt$ for the standard Riemann integral. We could also write this as

$$\int_a^b f(t)dt \quad = \quad \int_{[a,b]} f(t)dt$$

and we have learned that

$$\int_{[a,b]} f(t)dt \quad = \quad \int_{(a,b)} f(t)dt = \int_{(a,b]} f(t)dt = \int_{[a,b)} f(t)dt.$$

Note that we can thus say that we can compute $\int_E f(t)dt$ for $E \in \overline{\mathcal{B}}$ for sets E which are finite and have the form $[a, b]$, $(a, b]$, $[a, b)$ and (a, b). We can extend this easily to finite unions of disjoint intervals of the form E as given above by taking advantage of Theorem 4.5.3 to see

$$\int_{\cup_n E_n} f(t)dt \quad = \quad \sum_n \int_{E_n} f(t)dt.$$

However, the development of the Riemann integral is closely tied to the interval $[a, b]$ and so it is difficult to extend these integrals to arbitrary elements F of $\overline{\mathcal{B}}$. Still, we can see that the Riemann integral is defined on some subset of the sigma - algebra $\overline{\mathcal{B}}$.

From our discussions of the Riemann - Stieljes integral, we know that the Riemann integral can be interpreted as a Riemann - Stieljes integral with the integrator given by the identity function $id(x) = x$. Let's switch to a new notation. Define the function $\mu(x) = x$. Then for our allowable E, we can write $\int_E f(t)dt = \int_E f(t)d\mu(t)$ which we can further simplify to $\int_E f d\mu$ as usual. Note that μ is a function which assigns a real value which we interpret as *length* to all of the allowable sets E we have been discussing. In fact, note μ is a mapping which satisfies

(i): If E is the empty set, then the length of E is 0; i.e. $\mu(\emptyset) = 0$.

(ii): If E is the finite interval $[a, b]$, $(a, b]$, $[a, b)$ or (a, b), $\mu(E) = b - a$.

(iii): If (E_n) is a finite collection of disjoint intervals, then the length of the union is clearly the sum of the individual lengths; i.e. $\mu(\cup_n E_n) = \sum_n \mu(E_n)$.

However, μ is not defined on the entire sigma -algebra. Also, it seems that we would probably like to extend (iii) above to countable disjoint unions as it is easy to see how that would arise in practice. If we could find a way to extend the usual length calculation of an interval to the full sigma -algebra, we could then try to extend the notion of integration as well.

It turns out we can do all of these things but we can not do it by reusing our development process from Riemann integration. Instead, we must focus on developing a theory that can handle integrators which are mappings μ defined on a full sigma - algebra. It is time to precisely define what we mean by such a mapping.

Definition 10.0.1. Measures

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra of subsets in X. We say $\mu : \mathcal{S} \rightarrow \overline{\Re}$ is a measure on $\mathcal{S}$ if

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

(ii): $\mu(E) \geq 0$, *for all* $E \in \mathcal{S}$,

(iii): *μ is countably additive on $\mathcal{S}$; i.e. if $(E_n) \subseteq \mathcal{S}$ is a countable collection of disjoint sets, then $\mu(\cup_n E_n) = \sum_n \mu(E_n)$.*

We also say $(X, \mathcal{S}, \mu)$ is a measure space. If $\mu(X)$ is finite, we say μ is a finite measure. Also, even if $\mu(X) = \infty$,the measure μ is almost finite *if we can find a collection of measurable sets (F_n) so that $X = \cup_n F_n$ with $\mu(F_n)$ finite for all n. In this case, we say the measure μ is σ - finite.*

We can drop the requirement that the mapping μ be non negative. The resulting mapping is called a *charge* instead of a measure. This will be important later.

Definition 10.0.2. Charges

Let X be a nonempty set and $\mathcal{S}$ a sigma - algebra of subsets in X. We say $\nu : \mathcal{S} \rightarrow \Re$ is a charge on $\mathcal{S}$ if

(i): $\nu(\emptyset) = 0$,

(ii): *ν is countably additive on $\mathcal{S}$; i.e. if $(E_n) \subseteq \mathcal{S}$ is a countable collection of disjoint sets, then $\nu(\cup_n E_n) = \sum_n \nu(E_n)$.*

Note that we want the value of the charge to be finite on all members of $\mathcal{S}$ as otherwise we could potentially have trouble with subsets having value ∞ and $-\infty$ inside a given set. That would then lead to undefined $\infty - \infty$ operations.

Let's look at some examples:

Example 10.0.1. *Let X be any nonempty set and let the sigma - algebra be $\mathcal{S} = \mathcal{P}(X)$, the power set of X. Define μ_1 on $\mathcal{S}$ by $\mu_1(E) = 0$ for all E. Then μ_1 is a measure, albeit not very interesting! Another non interesting measure is defined by $\mu_2(E) = \infty$ if E is not empty and 0 if $E = \emptyset$.*

Example 10.0.2. *Let X be any set and again let $\mathcal{S} = \mathcal{P}(X)$. Pick any element p in X. Define μ by $\mu(E) = 0$ if $p \notin E$ and 1 if $p \in E$. Then μ is a measure.*

Example 10.0.3. *Let X be the counting numbers, $\boldsymbol{N}$, and $\mathcal{S} = \mathcal{P}(\boldsymbol{N})$. Define μ by $\mu(E)$ is the cardinality of E if E is a finite set and ∞ otherwise. Then μ is a measure called the counting measure. Note that $\boldsymbol{N} = \cup_n \{1, \ldots, n\}$ for all n and $\mu(\{1, \ldots, n\}) = n$, which implies μ is a σ - finite measure.*

Example 10.0.4. *This example is just a look ahead to future material we will be covering. Let $\overline{\mathcal{B}}$ be the extended Borel sigma - algebra. We will show later there is a measure $\lambda : \overline{\mathcal{B}} \rightarrow \overline{\Re}$ that extends the usual idea of the length of an interval. That is, if E is a finite interval of the form (a, b), $[a, b)$, $(a, b]$ or $[a, b]$, then the length of E is $b - a$ and $\lambda(E) = b - a$. Further, if the interval has infinite length, (for example, E is $(-\infty, a)$), then $\lambda(E) = \infty$ also. The measure λ will be called Borel measure and since $\overline{\Re} = \cup_n [-n, n]$, we see Borel measure is a σ - finite measure. The sets in $\overline{\mathcal{B}}$ are called Borel measurable sets.*

Example 10.0.5. *We will be able to show that there is a larger sigma - algebra $\mathcal{M}$ of subsets of $\overline{\Re}$ and a measure μ defined on $\mathcal{M}$ which also returns the usual length of intervals. Hence, $\overline{\mathcal{B}} \subseteq \mathcal{M}$ strictly (i.e. there are sets in $\mathcal{M}$ not in $\overline{\mathcal{B}}$) with $\mu = \lambda$ on $\overline{\mathcal{B}}$. This measure will be called Lebesgue measure and the sets in $\mathcal{M}$ will be called Lebesgue measurable sets. The proof that there are Lebesgue measurable sets that are not Borel sets will require a non constructive argument using the Axion of Choice. Further, we will be able to show that the Lebesgue sigma - algebra is not the entire power set as there are non Lebesgue measurable sets. The proof that such sets exist requires the use of the interesting functions built using Cantor sets discussed in Chapter 6.*

Example 10.0.6. *In the setting of Borel measure on $\overline{\Re}$, we will be able to show that if g is a continuous and monotone increasing function of $\Re$, then there is a measure, λ_g defined on $\overline{\mathcal{B}}$ which satisfies*

$$\lambda_g(E) = \int_E dg$$

for any finite interval E. Here, $\int_E dg$ is the usual Riemann - Stieljes integral.

10.1 Some Basic Properties Of Measures

Lemma 10.1.1. Monotonicity

Let $(X, \mathcal{S}, \mu)$ be a measure space. If $E, F \in \mathcal{S}$ with $E \subseteq F$, then $\mu(E) \leq \mu(F)$. Moreover, if $\mu(E)$ is finite, then $\mu(F \setminus E) = \mu(F) - \mu(E)$.

Proof. *We know $F = E \cup (F \setminus E)$ is a disjoint decomposition of F. By the countable additivity of μ, it follows immediately that $\mu(F) = \mu(E) + \mu(F \setminus E)$. Since μ is nonnegative, we see $\mu(F) \geq \mu(E)$. Finally, if $\mu(E)$ is finite, then subtraction is allowed in $\mu(F) = \mu(E) + \mu(F \setminus E)$ which leads to $\mu(F \setminus E) = \mu(F) - \mu(E)$.* ∎

Lemma 10.1.2. The Measure Of Monotonic Sequence Of Sets

Let $(X, \mathcal{S}, \mu)$ be a measure space.

(i): If (E_n) is an increasing sequence of sets in $\mathcal{S}$ (i.e. $E_n \subseteq E_{n+1}$ for all n), then $\mu(\cup_n E_n) = \lim_n \mu(E_n)$.

(ii): If (F_n) is an decreasing sequence of sets in $\mathcal{S}$ (i.e. $F_{n+1} \subseteq F_n$ for all n) and $\mu(F_1)$ is finite, then $\mu(\cap_n F_n) = \lim_n \mu(F_n)$.

Proof. *To prove (i), if there is an index n_0 where $\mu(E_{n_0}$ is infinite, then by the monotonicity of μ, we must have $\infty = \mu(E_{n_0} \leq \mu(\cup_n E_n)$. Hence, $\mu(\cup_n E_n) = \infty$. However, since $E_{n_0} \subseteq E_n$ for all $n \geq n_0$, again by monotonicity, $n \geq n_0$ implies $\mu(E_n) = \infty$. Thus, $\lim_n \mu(E_n) = \mu(\cup_n E_n) = \infty$. On the other hand, if $\mu(E_n)$ is finite for all n, define the disjoint sequence of set (A_n) as follows:*

$$\begin{array}{rcl} A_1 & = & E_1 \\ A_2 & = & E_2 \setminus E_1 \\ A_3 & = & E_3 \setminus E_2 \\ \vdots & \vdots & \vdots \\ A_n & = & E_n \setminus E_{n-1} \end{array}$$

We see $\cup_n A_n = \cup_n E_n$ and since μ is countably additive, we must have $\mu(\cup_n A_n) = \sum_n \mu(A_n)$. Since by assumption $\mu(E_n)$ is finite in this case, we know $\mu(A_n) = \mu(E_n) - \mu(E_{n-1})$. It follows that

$$\begin{aligned} \sum_{k=1}^{n} \mu(A_k) &= \mu(E_1) + \sum_{k=2}^{n} \Big(\mu(E_k) - \mu(E_{k-1}) \Big) \\ &= \mu(E_1) + \mu(E_n) - \mu(E_1) \\ &= \mu(E_n). \end{aligned}$$

We conclude

$$\begin{aligned} \mu(\cup_n E_n) &= \mu(\cup_n A_n) \\ &= \lim_n \sum_{k=1}^{n} \mu(A_k) \\ &= \lim_n \mu(E_n) \end{aligned}$$

this proves the validity of (i). Next, for (ii), construct the sequence of sets (E_n) by

$$\begin{array}{rcl} E_1 & = & \emptyset \\ E_2 & = & F_1 \setminus F_2 \\ E_3 & = & F_1 \setminus F_3 \\ \vdots & \vdots & \vdots \\ E_n & = & F_1 \setminus F_n. \end{array}$$

Then (E_n) is an increasing sequence of sets which are disjoint and so by (i), $\mu(\cup_n E_n) = \lim_n \mu(E_n)$. Since $\mu(F_1)$ is finite, we then know that $\mu(E_n) = \mu(F_1) - \mu(F_n)$. Hence, $\mu(\cup_n E_n) = \mu(F_1) - \lim_n \mu(F_n)$. Next, note by De Morgan's Laws,

$$\begin{aligned}
\mu(\cup_n E_n) &= \mu(\cup_n F_1 \cap F_n^C) \\
&= \mu(F_1 \cap \cup_n F_n^C) \\
&= \mu(F_1 \cap \Big(\cap_n F_n\Big)^C) \\
&= \mu(F_1 \setminus \Big(\cap_n F_n\Big)).
\end{aligned}$$

Thus, since $\mu(F_1)$ is finite and $\cap_n F_n \subseteq F_1$, we have $\mu(\cup_n E_n) = \mu(F_1) - \mu(\cap_n F_n)$. Combining these results, we have

$$\mu(F_1) - \lim_n \mu(F_n) = \mu(F_1) - \mu(\cap_n F_n).$$

The result then follows by canceling $\mu(F_1)$ from both sides which is allowed as this is a finite number. ■

We will now develop a series of ideas involving sequences of sets.

Definition 10.1.1. Limit Inferior And Superior Of Sequences Of Sets

Let X be a nonempty set and (A_n) be a sequence of subsets of X. The limit inferior of (A_n) is defined to be the set

$$\liminf = \underline{\lim}(A_n) = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n$$

while the limit superior of (A_n) is defined by

$$\limsup = \overline{\lim}(A_n) = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

It is convenient to have a better characterization of these sets.

Lemma 10.1.3. Characterizing Limit Inferior And Superiors Of Sequences Of Sets

Let (A_n) be a sequence of subsets of the nonempty set X. Then we have

$$\liminf(A_n) = \{x \in X \mid x \in A_k \textit{ for all but finitely many indices } k\} = B$$

and

$$\limsup(A_n) = \{x \in X \mid x \in A_k \textit{ for infinitely many indices } k\} = C$$

Proof. *We will prove the statement about $\liminf(A_n)$ first. Let $x \in B$. If there are no indices k so that $x \notin A_k$, then $x \in \cap_{n=1}^{\infty}$ telling us that $x \in \liminf(A_n)$. On the other hand, if there are a finite number of indices k that satisfy $x \notin A_k$, we can label these indices as $\{k_1, \ldots, k_p\}$ for some positive integer p. Let*

k^* be the maximum index in this finite list. Then, if $k > k^*$, $x \in \cap_{n=k}^{\infty}$. This implies immediately that $x \in \liminf(A_n)$. Conversely, if $x \in \liminf(A_n)$, there is an index k_0 so that $x \in \cap_{n=k_0}^{\infty}$. This implies that x can fail to be in at most a finite number of A_k where $k < k_0$. Hence, $x \in B$.

Next, we prove that $\limsup(A_n) = C$. If $x \in C$, then if there were an index m_0 so that $x \notin \cup_{n=m_0}^{\infty}$, then x would belong to only a finite number of sets A_k which contradicts the definition of the set C. Hence, there is no such index m_0 and so $x \in \cup_{n=m}^{\infty}$ for all m. This implies $x \in \limsup(A_n)$. On the other hand, if $x \in \limsup(A_n)$, then $x \in \cup_{n=m}^{\infty}$ for all m. So, if x was only in a finite number of sets A_n, there would be a largest index m^* satisfying $x \in A_{m^*}$ but $x \notin A_m$ if $m > m^*$. But this then says $x \notin \limsup(A_n)$. This is a contradiction. Thus, our assumption that x was only in a finite number of sets A_n is false. This implies $x \in C$. ■

Lemma 10.1.4. Limit Inferiors And Superiors Of Monotone Sequences Of Sets

Let X be a nonempty set. Then

(i): If (A_n) is an increasing sequence of subsets of X, then

$$\liminf(A_n) = \limsup(A_n) = \bigcup_{n=1}^{\infty} A_n.$$

(ii): If (A_n) is a decreasing sequence of subsets of X, then

$$\liminf(A_n) = \limsup(A_n) = \bigcap_{n=1}^{\infty} A_n.$$

(iii): If (A_n) is an arbitrary sequence of subsets of X, then

$$\emptyset \subseteq \liminf(A_n) \subseteq \limsup(A_n)$$

Proof.

(i): If $x \in \limsup(A_n)$, then $x \in \cup_{n=1}^{\infty} A_n$. Conversely, if $x \in \cup_{n=1}^{\infty} A_n$, there is an index n_0 so that $x \in A_{n_0}$. But since the sequence (A_n) is increasing, this means $x \in A_n$ for all $n > n_0$ also. Hence, $x \in \cup_{n=m}^{\infty} A_n$ for all indices $m \geq n_0$. However, it is also clear that x is in any union that starts at n smaller than n_0. Thus, x must be in $\cap_{m=1}^{\infty} \cup_{n=m}^{\infty} A_n$. But this is the set $\limsup(A_n)$. We conclude $\limsup(A_n) = \cup_{n=1}^{\infty}$. Now look at the definition of $\liminf(A_n)$. Since A_n is monotone increasing, $\cap_{n=m}^{\infty} A_n = A_m$. Hence, it is immediate that $\liminf(A_n) = \cup_{n=1}^{\infty}$.

(ii): the argument for this case is similar to the argument for case (i) and is left to you.

(iii): it suffices to show that $\liminf(A_n) \subseteq \limsup(A_n)$. If $x \in \liminf(A_n)$, by Lemma 10.1.3, x belongs to all but finitely many A_n. Hence, x belongs to infinitely many A_n. Then, applying Lemma 10.1.3 again, we have the result. ■

There will be times when it will be convenient to write an arbitrary union of sets as a countable union of disjoint sets. In the next result, we show how this is done.

Lemma 10.1.5. Disjoint Decompositions Of Unions

Let X be a nonempty set and let (A_n) be a sequence of subsets of X. Then there exists a sequence of mutually disjoint set (F_n) satisfying $\cup_n A_n = \cup_n F_n$.

Proof. *Define sets E_n and F_n as follows:*

$$\begin{aligned}
E_0 &= \emptyset, \quad F_1 = A_1 \setminus E_0 = A_1 \\
E_1 &= A_1, \quad F_2 = A_2 \setminus E_1 = A_2 \setminus A_1 \\
E_2 &= A_1 \bigcup A_2, \quad F_3 = A_3 \setminus E_2 = A_3 \setminus \Big(A_1 \bigcup A_2\Big) \\
E_3 &= \bigcup_{k=1}^{3} A_k, \quad F_4 = A_4 \setminus E_3 = A_4 \setminus \Big(\bigcup_{k=1}^{3} A_k\Big) \\
\vdots &= \vdots, \quad \vdots \\
E_n &= \bigcup_{k=1}^{n} A_k, \quad F_{n+1} = A_{n+1} \setminus E_n = A_4 \setminus \Big(\bigcup_{k=1}^{n} A_k\Big)
\end{aligned}$$

Note that (E_n) forms a monotonically increasing sequence of sets with cup$_n A_n$ $\cup_n E_n$. We claim the sets F_n are mutually disjoint and $\cup_{j=1}^{n} f_j = \cup_{j=1}^{n} A_j$. We do this by induction.

Subproof. *Basis: It is clear that F_1 and F_2 are disjoint and $F_1 \cup F_2 = A_1 \cup A_2$. Induction: We assume that (F_k) are mutually disjoint for $1 \leq k \leq n$ and $\cup_{j=1}^{k} f_j = \cup_{j=1}^{k} A_j$ for $1 \leq k \leq n$ as well. Then*

$$\begin{aligned}
F_{n+1} &= A_{n+1} \setminus E_n \\
&= A_{n+1} \bigcap \Big(\bigcup_{j=1}^{n} A_j\Big)^C \\
&= \bigcap_{j=1}^{n} \Big(A_{n+1} \bigcap A_j^C\Big).
\end{aligned}$$

Now, by construction, $F_j \subseteq A_j$ for all j. However, from the above expansion of F_{n+1}, we see $F_{n+1} \subseteq A_j^C$ for all $1 \leq j \leq n$. This tells us $F_{n+1} \subseteq F_j^C$ for these indices also. We conclude F_{n+1} is disjoint from all the previous F_j. This shows (F_j) is a collection of mutually disjoint sets for $1 \leq j \leq n+1$. This proves the first part of the assertion. To prove the last part, note

$$\begin{aligned}
\bigcup_{j=1}^{n+1} F_j &= \bigcup_{j=1}^{n} F_j \bigcup F_{n+1} \\
&= \bigcup_{j=1}^{n} A_j \bigcup \Big(A_{n+1} \setminus \Big(\bigcup_{j=1}^{n} A_j\Big)\Big) \\
&= \bigcup_{j=1}^{n+1} A_j.
\end{aligned}$$

This completes the induction step. We conclude that this proposition holds for all n. □

Since the claim holds, it is then obvious that $\cup_{j=1}^{n} f_j = \cup_{j=1}^{n} A_j$. ■

To finish this section on measures, we want to discuss the idea that a property holds except on a set of measure zero. Recall, this subject came up when we discussed the content of a subset of $\Re$ earlier in Section 5.3. However, we can extend this concept of an arbitrary measure space $(X, \mathcal{S}, \mu)$ as follows.

Definition 10.1.2. Propositions Holding Almost Everywhere

Let $(X, \mathcal{S}, \mu)$ be a measure space. We say a proposition $\mathcal{P}$ holds almost everywhere on X if $\{x \in X \mid \mathcal{P} \text{ does not hold }\}$ has μ measure zero. We usually say the proposition holds μ a.e. rather than writing out the phrase μ almost everywhere. *Also, if the measure μ is understood from context, we usually just say the proposition hold a.e. to make it even easier to write down.*

Comment 10.1.1. *Given the measure space $(X, \mathcal{S}, \mu)$, if f and g are extended real valued functions on X which are measurable, we would say $f = g$ μ a.e. if $\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0$.*

Comment 10.1.2. *Given the measure space $(X, \mathcal{S}, \mu)$, If (f_n) is a sequence of measurable extended real valued functions on the X, and $f : X \to \overline{\Re}$ is another measurable function on X, we would say f_n converges pointwise a.e. to f if the set $\{x \in X \mid f_n(x) \not\to f(x)\}$ has measure 0. We would usually write $f_n \to f$ pointwise μ a.e.*

10.2 Integration

In this section, we will introduce an abstract notion of integration on the measure space $(X, \mathcal{S}, \mu)$. Recall that $M(X, \mathcal{S})$ denotes the class of extended real valued measurable functions f on X. First we introduce a standard notation for some useful classes of functions. When we want to restrict our attention to the non negative members of $M(X, \mathcal{S})$, we will use the notation that $f \in M^+(X, \mathcal{S})$.

To construct an abstract integration process on the measure space $(X, \mathcal{S}, \mu)$, we begin by defining the *integral* of a class of functions which can be used to approximate any function f in $M^+(X, \mathcal{S})$.

Definition 10.2.1. Simple Functions

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $f : X \to \Re$ be a function. We say f is a simple function if the range of f is a finite set and f is $\mathcal{S}$ measurable. This implies the following standard unique representation of f. Since the range is finite, there is an positive integer N and distinct numbers a_j, $1 \leq j \leq N$ so that

(i): the sets $E_j = f^{-1}(a_j)$ are measurable and mutually disjoint for $1 \leq j \leq N$,

(ii): $X = \cup_{j=1}^{N} E_j$,

(iii): f has the characterization

$$f(x) \quad = \quad \sum_{j=1}^{N} a_j I_{E_j}(x).$$

We then define the *integral* of a simple function as follows.

Definition 10.2.2. The Integral Of A Simple Function

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $\phi : X \to \Re$ be a simple function. Let

$$\phi(x) = \sum_{j=1}^{N} a_j I_{E_j}(x),$$

be the standard representation of ϕ where the numbers a_j are distinct and the sets E_j are mutually disjoint, cover X, and are measurable for $1 \leq j \leq N$ for some positive integer N. Then the integral of ϕ with respect to the measure μ is the extended real valued number

$$\int \phi \, d\mu = \sum_{j=1}^{N} a_j \, \mu(E_j).$$

Comment 10.2.1. *We note that $\int \phi d\mu$ can be $+\infty$. Recall, our convention that $0 \cdot \infty = 0$. Hence, if one of the values a_j is 0, the contribution to the integral is $0\mu(E_j)$ which is 0 even if $\mu(E_j) = \infty$. Further, note the 0 function on X can be defined as $I_\emptyset$ which is a simple function. Hence, $\int 0 \, d\mu = 0$.*

Using this, we can define the integral of any function in $M^+(X, \mathcal{S})$.

Definition 10.2.3. The Integral Of A Nonnegative Measurable Function

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $f \in M^+(X, \mathcal{S}), \mu)$. For convenience of notation, let $\mathcal{F}^+$ denote the collection of all non negative simple functions on X. Then, the integral of f with respect to the measure μ is the extended value real number

$$\int f \, d\mu = \sup\{\int \phi \, d\mu \,|\, \phi \in \mathcal{F}^+, \, \phi \leq f\}.$$

If $E \in \mathcal{S}$, we define the integral of f over E with respect to μ to be

$$\int_E f \, d\mu = \int f I_E \, d\mu.$$

It is time to prove some results about this new abstract version of integration.

Lemma 10.2.1. Properties Of Simple Function Integrations

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $\phi, \psi \in M^+(X, \mathcal{S}))$ be simple functions. Then,

(i): If $c \geq 0$ is a real number, then $c\phi$ is also a simple function and $\int c\phi \, d\mu = c \int \phi \, d\mu$.

(ii): $\phi + \psi$ is also a simple function and $\int (\phi + \psi) \, d\mu = \int \phi \, d\mu + \int \psi \, d\mu$.

(iii): The mapping $\lambda : \mathcal{S} \to \Re$ defined by $\lambda(E) = \int_E \phi \, d\mu$ for all E in $\mathcal{S}$ is a measure.

Proof. *Let ϕ have the standard representation*

$$\phi(x) = \sum_{j=1}^{N} a_j I_{E_j}(x),$$

where the numbers a_j are distinct, the sets E_j are mutually disjoint, cover X, and are measurable for $1 \leq j \leq N$ for some positive integer N. Similarly, let ψ have the standard representation

$$\psi(x) = \sum_{k=1}^{M} b_k I_{F_k}(x),$$

where the numbers b_k are distinct, the sets F_k are mutually disjoint, cover X, and are measurable for $1 \leq k \leq M$ for some positive integer M. Now to the proofs of the assertions: (i):
First, if $c = 0$, $c\phi = 0$ and $\int 0 d\mu = 0 \int \phi d\mu$. Next, if $c > 0$, then it is easy to see $c\phi$ is a simple function with representation

$$c\phi(x) = \sum_{j=1}^{N} c a_j I_{E_j}(x),$$

and hence, by the definition of the integral of a simple function

$$\begin{aligned} \int c\phi \, d\mu &= \sum_{j=1}^{N} c a_j \, \mu(E_j) \\ &= c\left(\sum_{j=1}^{N} a_j \, \mu(E_j)\right) \\ &= c \int \phi \, d\mu. \end{aligned}$$

(ii):
This one is more interesting to prove. First, to prove $\phi + \psi$ is a simple function, all we have to do is find its standard representation. From the standard representations of ϕ and ψ, it is clear the sets $F_k \cap E_j$ are mutually disjoint and since $X = \cup E_j = \cup F_k$, we have the identities

$$F_k = \bigcup_{j=1}^{N} F_k \cap E_j, \quad \textit{and} \quad E_j = \bigcup_{k=1}^{M} F_k \cap E_j.$$

Now define $h : X \to \Re$ by

$$h(x) = \sum_{j=1}^{N} \sum_{k=1}^{M} (a_j + b_k) \, I_{F_k \cap E_j} \, (x).$$

Next, since $X = \cup_j \cup_k F_k \cap E_j$, given $x \in X$, there are indices k_0 and j_0 so that $x \in F_{k_0} \cap E_{j_0}$. Thus,

$$\phi(x) + \psi(x) = a_{j_0} \, I_{E_{j_0}} + b_{k_0} \, I_{F_{k_0}} = a_{j_0} + b_{k_0}.$$

From the above argument, we see $h(x) = \phi(x) + \psi(x)$ for all x in X. It follows that the range of h is finite and hence it is a measurable simple function, but we still do not know its standard representation.

To find the standard representation, let c_i, $1 \leq i \leq P$ be the set of distinct numbers formed by the collection $\{a_j + b_k \mid 1 \leq j \leq N, \, 1 \leq k \leq M\}$. Then let U_i be the set of index pairs (j, k) that satisfy

$c_i = a_j + b_k$. *Finally, let*

$$G_i = \bigcup_{(j,k)\in U_i} E_j \cap F_k.$$

Since the sets $F_k \cap E_j$ *are mutually disjoint, we have*

$$\mu(G_i) = \sum_{(j,k)\in U_i} \mu(E_j \cap F_k).$$

It follows that

$$h(x) = \sum_{i=1}^{P} c_i \, I_{G_i}$$

is the standard representation of $h = \phi + \psi$. *Thus*

$$\begin{aligned}
\int h \, d\mu &= \int (\phi + \psi) \, d\mu \; = \sum_{i=1}^{P} c_i \, \mu(G_i) \\
&= \sum_{i=1}^{P} c_i \left(\sum_{(j,k)\in U_i} \mu(E_j \cap F_k) \right) \\
&= \sum_{i=1}^{P} \sum_{(j,k)\in U_i} c_i \, \mu(E_j \cap F_k).
\end{aligned}$$

But we know that

$$\sum_{j=1}^{N} \sum_{k=1}^{M} = \sum_{i=1}^{P} \sum_{(j,k)\in U_i}.$$

Hence, we can write

$$\begin{aligned}
\int (\phi + \psi) \, d\mu &= \sum_{j=1}^{N} \sum_{k=1}^{M} (a_j + b_k) \, \mu(E_j \cap F_k) \\
&= \sum_{j=1}^{N} \sum_{k=1}^{M} a_j \, \mu(E_j \cap F_k) \; + \sum_{j=1}^{N} \sum_{k=1}^{M} b_k \, \mu(E_j \cap F_k).
\end{aligned}$$

This can be reorganized as

$$\begin{aligned}
\sum_{j=1}^{N} a_j \sum_{k=1}^{M} \mu(E_j \cap F_k) \; + \sum_{k=1}^{M} b_k \sum_{j=1}^{N} \mu(E_j \cap F_k) &= \sum_{j=1}^{N} a_j \, \mu(\bigcup_{k=1}^{M} E_j \cap F_k) \; + \sum_{k=1}^{M} b_k \, \mu(\bigcup_{j=1}^{N} E_j \cap F_k) \\
&= \sum_{j=1}^{N} a_j \, \mu(E_j) \; + \sum_{k=1}^{M} b_k \, \mu(F_k) \\
&= \int \phi \, d\mu \; + \int \psi \, d\mu.
\end{aligned}$$

(iii):
Given

$$\phi(x) \;=\; \sum_{j=1}^{N} a_j I_{E_j}(x),$$

it is easy to see that

$$\phi\, I_E(x) \;=\; \sum_{j=1}^{N} a_j I_{E \cap E_j}(x).$$

Further, it is straightforward to show that the mappings $\mu_j : (S) \to \Re$ *defined by* $\mu_j(A) = \mu(A \cap E_j)$ *for all* A *in* $\mathcal{S}$ *are measures on the sigma - algebras* $\mathcal{S} \cap E_j$ *for each* $1 \le j \le N$. *It is also easy to see that the finite linear combination of these measures given by* $\xi = \sum_{j=1}^{N} a_j\, \mu_j$ *is a measure on* $\mathcal{S}$ *itself. Thus, applying part (ii) of this lemma, we see*

$$\begin{aligned}
\lambda(E) &= \int \phi I_E \, d\mu \;=\; \int \phi I_{\cup_{j=1}^{N} E \cap E_j} \, d\mu \\
&= \int \left(\sum_{j=1}^{N} \phi I_{E \cap E_j} \right) d\mu \;=\; \sum_{j=1}^{N} \int a_j I_{E \cap E_j} \, d\mu \\
&= \sum_{j=1}^{N} a_j \mu(E \cap E_j) \;=\; \sum_{j=1}^{N} a_j\, \mu_j(E) \;=\; \xi(E).
\end{aligned}$$

We conclude $\lambda = \xi$ *and* λ *is a measure on* $\mathcal{S}$. ∎

Lemma 10.2.2. Monotonicity Of The Abstract Integral For Non Negative Functions
Let $(X, \mathcal{S}, \mu)$ *be a measure space and let* f *and* g *be in* $M^+(X, \mathcal{S})$ *with* $f \le g$. *Then,* $\int f d\mu \le \int g d\mu$. *Further, if* $E \subseteq F$ *with* E *and* F *measurable sets, then* $\int_E f d\mu \le \int_F f d\mu$.

Proof. *Let* ϕ *be a positive simple function which is dominated by* f*; i.e.,* $\phi \le f$. *Then* ϕ *is also dominated by* g *and so by the definition of the integral of* f*, we have*

$$\begin{aligned}
\int f d\mu &= \sup \{ \int \phi d\mu \,|\, 0 \le \phi \le f \} \\
&\le \sup \{ \int \psi d\mu \,|\, 0 \le \psi \le g \} \\
&= \int g d\mu.
\end{aligned}$$

Next, if $E \subseteq F$ *with* E *and* F *measurable sets, then* $f I_E \le f I_F$ *and from the first result, we have*

$$\int f I_E d\mu \;\le\; \int f I_F d\mu,$$

which implies the result we seek. ∎

10.3 Complete Measures And Equality a.e.

We know that is a sequence of extended real - valued measurable functions (f_n) converges pointwise to a function f, then the limit function is also measurable. But what if the convergence was pointwise a.e? Is it still true that the limit function is also measurable. In general, the answer is no. We have to add an additional property to the measure. We will motivate this with an example that we are not really fully prepared for, but it should make sense anyway.

Let $\boldsymbol{B} \cap [0,1]$ denote the Borel sigma - algebra of subsets of $[0,1]$. We will be able to show in later chapters, that there is a measure called Lebesgue measure, μ_L, defined on a sigma - algebra of subsets $\boldsymbol{L}$, the Lebesgue sigma - algebra, which extends the usual meaning of length in the following sense. If $[a,b]$ is a finite interval then the length of $[a,b]$ is the finite number $b-a$. Denote this length by $\ell([a,b])$. Then we can show that

$$\mu_L([a,b]) \quad = \quad \ell([a,b]) \; = \; b-a.$$

We can show also that every subset in $\boldsymbol{B}$ is also in $\boldsymbol{L}$. The restriction of μ_L to $\boldsymbol{B}$ is called Borel measure and we will denote it by μ_B.

We can argue that the Borel sigma - algebra is strictly contained in the Lebesgue sigma - algebra by using the special functions we constructed in Chapter 6. Recall that if $\boldsymbol{C}$ is a Cantor set constructed from the generating sequence (a_n) where $\lim 2^n a_n = 0$, we could show the content of $\boldsymbol{C}$ was 0. Then if we let Ψ be the mapping discussed above for this $\boldsymbol{C}$ in Section 6.3, we define the mapping mapping $g : [0,1] \to [0,1]$ by $g(x) = (\Psi(x) + x)/2$. The mapping g is quite nice: it is $1-1$, onto, strictly increasing and continuous. We also showed in the exercises in Section 6.3 that $g(\boldsymbol{C})$ is another Cantor set with $\lim 2^n a'_n = 1/2$, where (a'_n) is the generating sequence for $g(\boldsymbol{C})$.

Now it turns out that the notion of content and Lebesgue measure coincide. Thus, we can say since $\boldsymbol{C}$ is a Borel set,

$$\mu_B(\boldsymbol{C}) \quad = \quad \mu_L(\boldsymbol{C}) \; = \; 0.$$

Also, we can show that since $\lim 2^n a'_n = 1/2$,

$$\mu_B(g(\boldsymbol{C})) \quad = \quad \mu_L(g(\boldsymbol{C})) \; = \; 1/2.$$

A nonconstructive argument we will present later using the Axiom of Choice allows us to show that any Lebesgue measurable set with positive Lebesgue measure must contain a subset which is not in the Lebesgue sigma - algebra. So since $\mu_L(g(\boldsymbol{C})) = 1/2$, there is a set $F \subseteq g(\boldsymbol{C})$ which is not is $\boldsymbol{L}$. Thus, $g^{-1}(F) \subseteq \boldsymbol{C}$ which has Lebesgue measure 0. Lebesgue measure is a measure which has the property that every subset of a set of measure 0 must be in the Lebesgue sigma - algebra. Then, using the monotonicity of μ_L, we have $\mu_L(g^{-1}(F))$ is also 0. From the above remarks, we can infer something remarkable.

Let the mapping h be defined to be g^{-1}. Then h is also continuous and hence it is measurable with respect to the Borel sigma-algebra. Note since $\boldsymbol{B} \subseteq \boldsymbol{L}$, this tells us immediately that h is also measurable with respect to the Lebesgue sigma - algebra. Thus, $h^{-1}(U)$ is in the Borel sigma - algebra for all Borel sets U. But we know $h^{-1} = g$, so this tells us $g(U)$ is in the Borel sigma -algebra if U is a Borel set. Hence, if we chose $U = g^{-1}(F)$, then $g(U) = F$ would have to be a Borel set if U is a Borel set. However, we know that F is not in $\boldsymbol{L}$ and so it is also not a Borel set. We can only conclude that

$g^{-1}(F)$ can not be a Borel set. However, $g^{-1}(F)$ is in the Lebesgue sigma - algebra. Thus, there are Lebesgue measurable sets which are not Borel! Thus, the Borel sigma - algebra is strictly contained in the Lebesgue sigma - algebra!

We can use this example to construct another remarkable thing.

Comment 10.3.1. *Using all the notations from above, note the indicator function of $\boldsymbol{C}^C$, the complement of $\boldsymbol{C}$, is defined by*

$$I_{\boldsymbol{C}^C}(x) = \begin{cases} 1 & x \in \boldsymbol{C}^C \\ 0 & x \in \boldsymbol{C}. \end{cases}$$

We see $f = I_{C^C}$ is Borel measurable. Next, define a new mapping like this:

$$\phi(x) = \begin{cases} 1 & x \in \boldsymbol{C}^C \\ 2 & x \in \boldsymbol{C} \setminus g^{-1}(F) \\ 3 & x \in g^{-1}(F). \end{cases}$$

Note that $\phi = f$ a.e. with respect to Borel measure. However, ϕ is **not** *Borel measurable because $\phi^{-1}(3)$ is the set $g^{-1}(F)$ which is not a Borel set.*

We conclude that in this case, even though the two functions were equal a.e. with respect to Borel measure, only one was measurable! The reason this happens is that even though $\boldsymbol{C}$ has Borel measure 0, there are subsets of $\boldsymbol{C}$ which are not Borel sets!

Hence, in some situations, we will have to stipulate that the measure we are working with has the property that every subset of a set of measure zero is measurable. We make this formal with a definition.

Definition 10.3.1. Complete Measure

Let X be a nonempty set and $(X, \mathcal{S}, \mu)$ be a measure space. If $E \in \mathcal{S}$ with $\mu(E) = 0$ and $F \subseteq E$ implies $F \in \mathcal{S}$, we say μ is a complete measure. Further, it follows immediately that since $\mu(F) \leq \mu(E) = 0$, that $\mu(F) = 0$ also.

Comment 10.3.2. *This example above can be used in another way. Consider the composition of the measurable function I_C and the function g defined above. For convenience, let $W = g^{-1}(F)$ which is Lebesgue measurable. Then I_W is a measurable function. Consider*

$$\Big(I_W \circ g^{-1}\Big)(x) = \begin{cases} 1 & g^{-1}(x) \in W \\ 0 & g^{-1}(x) \in W^C \end{cases} = \begin{cases} 1 & x \in g(W) \\ 0 & x \in g(W^C) \end{cases} = \begin{cases} 1 & x \in F \\ 0 & x \in F^C \end{cases} = I_F.$$

But I_F is not a measurable function as F is not a measurable set! Hence, the composition of the measurable function I_W and the continuous function g^{-1} is not measurable. This is why we can only prove measurability with the order of the composition reversed as we did in Lemma 9.6.7.

Theorem 10.3.1. Equality a.e. Implies Measurability If The Measure Is Complete

Let X be a nonempty set and $(X, \mathcal{S}, \mu)$ be a measure space. Let f and g both be extended real valued functions on X with $f = g$ a.e. Then, if f is measurable, so is g.

Proof. *Let G be open in $\overline{\Re}$ and let $E = (f(x) \neq g(x))$. Then, by assumption, E is measurable and $\mu(E) = 0$. Then, we claim*

$$g^{-1}(G) = \Big(g^{-1}(G) \cap E\Big) \cup \Big(f^{-1}(G) \setminus E\Big).$$

If x is in $g^{-1}(G)$, then $g(x)$ is in $G \cap E$ or it is in $G \cap E^C$. Now if $g(x) \in E$, $g(x) \neq f(x)$, but if $g(x)$ is in the complement of E, g and f must match. Thus, we see x is in the right hand side. Conversely, if x is in $g^{-1}(G) \cap E$, x is clearly in $g^{-1}(G)$. Finally, if x is in $f^{-1}(G) \setminus E$, then since x is not in E, $f(x) = g(x)$. Thus, $x \in g^{-1}(G)$ also. We conclude $x \in g^{-1}(G)$. This shows the right hand side is contained in the left hand side. Combining these arguments, we conclude the two sets must be equal.

Since $g^{-1}(G) \cap E$ is a subset of E, the completeness of μ implies that $g^{-1}(G) \cap E$ is measurable and has measure 0. The measurability of f tells us that $f^{-1}(G) \setminus E$ is also measurable. Hence, $g^{-1}(G)$ is measurable implying g is measurable. ∎

If the measure μ is not complete, we can still prove the following.

Theorem 10.3.2. Equality a.e. Can Imply Measurability Even If The Measure Is Not Complete

Let X be a nonempty set and $(X, \mathcal{S}, \mu)$ be a measure space. Let f and g both be extended real valued functions on X with $f = g$ on the measurable set E^C with $\mu(E) = 0$. Then, if f is measurable and g is constant on E, g is measurable.

Proof. *We will repeat the notation of the previous theorem's proof. As before, if G is open, we can write*

$$g^{-1}(G) = \Big(g^{-1}(G) \cap E\Big) \cup \Big(f^{-1}(G) \setminus E\Big).$$

Then, since g is constant on E with value say c, we have

$$g^{-1}(G) = \left(\begin{cases} E & c \in G \\ \emptyset & c \notin G \end{cases}\right) \bigcup \Big(f^{-1}(G) \setminus E\Big) = \begin{cases} E \cup \Big(f^{-1}(G) \setminus E\Big) & c \in G \\ \Big(f^{-1}(G) \setminus E\Big) & c \notin G. \end{cases}$$

In both cases, the resulting set is measurable. Hence, we conclude g is measurable. ∎

Comment 10.3.3. *In Comment 10.3.1, we set*

$$\phi(x) = \begin{cases} 1 & x \in \boldsymbol{C}^C \\ 2 & x \in \boldsymbol{C} \setminus g^{-1}(F) \\ 3 & x \in g^{-1}(F). \end{cases}$$

and since ϕ was not constant on $E = \boldsymbol{C}$, ϕ was not measurable. However, if we had defined

$$\phi(x) = \begin{cases} 1 & x \in \boldsymbol{C}^C \\ c & x \in \boldsymbol{C}, \end{cases}$$

then ϕ would have been measurable!

10.4 Convergence Theorems

We are now ready to look at various types of interchange theorems for abstract integrals. We will be able to generalize the results of Chapter 5 substantially. There are three basic results: (i) The Monotone Convergence Theorem, (ii) Fatou's Lemma and (iii) The Lebesgue Dominated Convergence Theorem. We will examine each in turn.

Theorem 10.4.1. The Monotone Convergence Theorem
Let $(X, \mathcal{S}, \mu)$ be a measure space and let (f_n) be an increasing sequence of functions in $M^+(X, \mathcal{S})$. Let $f : X \to \overline{\Re}$ be an extended real valued function such that $f_n \to f$ pointwise on X. Then f is also in $M^+(X, \mathcal{S})$ and

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

Proof. *Since f_n converges to f pointwise, we know that f is measurable by Theorem 9.6.5. Further, since $f_n \geq 0$ for all n on X, it is clear that $f \geq 0$ also. Thus, $f \in M^+(X, \mathcal{S})$. Since $f_n \leq f_{n+1} \leq f$, the monotonicity of the integral tells us*

$$\int f_n d\mu \leq \int f_{n+1} d\mu \leq \int f d\mu.$$

Hence, $\int f_n d\mu$ is an increasing sequence of real numbers bounded above by $\int f d\mu$. Of course, this limit could be ∞. Thus, we have the inequality

$$\int f_n d\mu \leq \int f d\mu.$$

We no show the reverse inequality, $\int f d\mu \leq \int f_n d\mu$. Let α be in $(0, 1)$ and choose any non negative simple function ϕ which is dominated by f. Let

$$A_n \;=\; \{x \,|\, f_n(x) \geq \alpha\, \phi(x)\}.$$

We claim that $X = \cup_n A_n$. If this was not true, then there would be an x which is not in any A_n. This implies x is in $\cap_n A_n^C$. Thus, using the definition of A_n, $f_n(x) < \alpha\, \phi(x)$ for all n. Since f_n is increasing and converges pointwise to f, this tells us

$$f(x) \;\leq\; \alpha\, \phi(x) \;\leq\; \alpha\, \phi(x).$$

We can rewrite this as $(1 - \alpha)\, f(x) \leq 0$ and since $1 - \alpha$ is positive by assumption, we can conclude $f(x) \leq 0$. But f is non negative, so combining, we see $f(x) = 0$. Since f dominates ϕ, we must have $\phi(x) = 0$ too. However, if this is true, $f_n(x)$ must be 0 also. Hence, $f_n(x) = 0 \geq \alpha\phi(x) = 0$ for all n. This says $x \in A_n$ for all n. This is a contradiction; thus, $X = \cup_n A_n$.

Next, since f and $\alpha\phi$ are measurable, so is $f - \alpha\phi$. This implies $\{x \,|\, f(x) - \alpha\phi(x) \geq 0\}$ is a measurable set. Therefore, A_n is measurable for all n. Further, it is easy to $A_n \subseteq A_{n+1}$ for all n; hence, (A_n) is an increasing sequence of measurable sets. Then, we know by the monotonicity of the integral, that

$$\int_{A_n} \alpha\phi d\mu \leq \int_{A_n} f_n d\mu \leq \int f d\mu.$$

Next, we know that $\lambda(E) = \int_E \phi d\mu$ *defines a measure. Thus,*

$$\lim_n \lambda(A_n) \;=\; \lambda(\cup_n A_n) \;=\; \lambda(X).$$

Replacing λ *by its meaning in terms of* ϕ*, we have*

$$\int \phi d\mu \;=\; \lim_n \int_{A_n} \phi d\mu.$$

Multiplying through by the positive number α*, we have*

$$\alpha \int \phi d\mu \;=\; \lim_n \int_{A_n} \alpha\, \phi d\mu \;\leq\; \lim_n \int_{A_n} \alpha\, f_n d\mu.$$

Thus, for all $\alpha \in (0,1)$*, we have*

$$\alpha \int \phi d\mu \;\leq\; \lim_n \int_{A_n} \alpha\, f_n d\mu.$$

Letting $\alpha \rightarrow 1$*, we obtain*

$$\int \phi d\mu \;\leq\; \lim_n \int_{A_n} \alpha\, f_n d\mu.$$

Since the above inequality is valid for all non negative simple functions dominated by f*, we have immediately*

$$\int f d\mu \;\leq\; \lim_n \int_{A_n} \alpha\, f_n d\mu,$$

which provides the other inequality we need to prove the result. ■

This has an immediate extension to series of nonnegative functions.

Theorem 10.4.2. The Extended Monotone Convergence Theorem

Let $(X, \mathcal{S}, \mu)$ *be a measure space and let* (g_n) *be a sequence of functions in* $M^+(X, \mathcal{S})$*. Then, the sequence of partial sums,*

$$S_n \;=\; \sum_{k=1}^{n} g_n$$

converges pointwise on X *to the extended real valued non negative valued function* $S = \sum_{k=1}^{\infty} g_n$*. Further,* S *is also in* $M^+(X, \mathcal{S})$ *and*

$$\lim_n \int S_n d\mu\Big) \;=\; \int S\, d\mu.$$

This can also be written in series notation as

$$\sum_{k=1}^{\infty} \int f_k\, d\mu \;=\; \int \Big(\sum_{k=1}^{\infty} \int f_k\Big)\, d\mu.$$

Proof. *To prove this result, just apply the Monotone Convergence Theorem to the sequence of partial sums* (S_n)*.* ■

The Monotone Convergence Theorem allows us to prove that this notion of integration is additive and linear for positive constants.

Theorem 10.4.3. Abstract Integration Is Additive

Let $(X, \mathcal{S}, \mu)$ be a measure space and let f and g be functions in $M^+(X, \mathcal{S})$. Further, let α be a non negative real number. Then

(i): αf is in $M^+(X, \mathcal{S})$ and

$$\int \alpha f d\mu = \alpha \int f d\mu.$$

(ii): Also, $f + g$ is in $M^+(X, \mathcal{S})$ and

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

Proof.

(i): The case $\alpha = 0$ is clear, so we may assume without loss of generality that $\alpha > 0$. We know from Theorem 9.6.8 that there is a sequence of non negative simple functions (ϕ_n) which are increasing and converge up to f on X. Hence, since $\alpha > 0$, we also know that $\alpha \phi_n \uparrow \alpha f$. Thus, by the Monotone Convergence Theorem, αf is in $M^+(X, \mathcal{S})$ and

$$\lim_n \int \alpha \phi_n d\mu = \int \alpha f d\mu.$$

From Lemma 10.2.1, we know that $\int \alpha \phi_n d\mu = \alpha \int \phi_n d\mu$. Thus,

$$\int \alpha f d\mu = \alpha \lim_n \int \phi_n d\mu = \alpha \int f d\mu.$$

(ii): If we apply Theorem 9.6.8 to f and g, we find two sequences of increasing simple functions (ϕ_n) and $\psi_n)$ so that $\phi_n \uparrow f$ and $\psi_n \uparrow g$. Thus, $(\phi_n + \psi_n) \uparrow (f + g)$. Hence, by the Monotone Convergence Theorem, $f + g$ is in $M^+(X, \mathcal{S})$ and

$$\begin{aligned} \int (f+g) d\mu &= \lim_n \int (\phi_n + \psi_n) d\mu = \lim_n \int \phi_n d\mu + \lim_n \int \psi_n d\mu \\ &= \int f d\mu + \int g d\mu. \end{aligned}$$

■

Theorem 10.4.4. Fatou's Lemma

Let $(X, \mathcal{S}, \mu)$ be a measure space and let (f_n) be a sequence of functions in $M^+(X, \mathcal{S})$. Then

$$\int \liminf f_n \, d\mu \leq \int \liminf f_n \, d\mu.$$

Proof. *Recall*

$$\begin{aligned}\liminf f_n(x) &= \sup_m \inf_{n\geq m} f_n(x)\\ &= \lim_m \inf_{n\geq m} f_n(x),\\ \limsup f_n(x) &= \inf_m \sup_{n\geq m} f_n(x)\\ &= \lim_m \sup_{n\geq m} f_n(x).\end{aligned}$$

Further, if we define $g_m = \inf_{n\geq m} f_n(x)$*, we know*

$$g_m \uparrow \liminf f_n(x).$$

It follows immediately that g_m *is measurable for all* m *and by the monotonicity of the integral*

$$\int g_m\, d\mu \leq \int f_n(x)\, d\mu \;\; \forall\, n \geq m.$$

This implies that $\int g_m\, d\mu$ *is a lower bound for the set of numbers* $\{\int f_n(x)\, d\mu\}$ *and so by definition of the infimum,*

$$\int g_m\, d\mu \leq \inf_{n\geq m}\left(\int f_n(x)\, d\mu\right).$$

Let α_m *denote the number* $\inf_{n\geq m}\left(\int f_n(x)\, d\mu\right)$. *Then,* $\alpha_m \uparrow \liminf \int f_n d\mu$. *We see*

$$\begin{aligned}\lim_m \int g_m\, d\mu &\leq \lim_m \inf_{n\geq m}\left(\int f_n(x)\, d\mu\right)\\ &= \lim_m \alpha_m = \liminf \int f_n d\mu.\end{aligned}$$

But since $g_m \uparrow \liminf f_n(x)$*, this implies*

$$\int \liminf f_n(x)\, d\mu \leq \liminf \int f_n\, d\mu.$$

∎

These results allow us to construct additional measures.

Theorem 10.4.5. Constructing Measures From Non Negative Measurable Functions

Let $(X, \mathcal{S}, \mu)$ be a measure space and let f be a function in $M^+(X, \mathcal{S})$. Then $\lambda : \mathcal{S} \to \overline{\Re}$ defined by

$$\lambda(E) = \int_E f \, d\mu, \; E \in \mathcal{S}$$

is a measure.

Proof. *If is clear $\lambda(\emptyset)$ is 0 and that $\lambda(E)$ is always non negative. To show that λ is countably additive, let (E_n) be a sequence of disjoint measurable sets in $\mathcal{S}$ and let $E = \cup_n E_n$, be their union. Then E is measurable. Define*

$$f_n = \sum_{k=1}^{n} f I_{E_k} = f I_{\cup_{k=1}^{n} E_k}.$$

We note that $f_n \uparrow f I_E$ and so by the Monotone Convergence Theorem,

$$\lambda(E) = \int_E f \, d\mu = \int f I_E d\mu = \lim_n \int f_n \, d\mu.$$

But,

$$\begin{aligned} \int f_n \, d\mu &= \int \left(f I_{E_n} \right) d\mu \\ &= \sum_{k=1}^{n} int \, f I_{E_k} \, d\mu = \sum_{k=1}^{n} \lambda(E_k). \end{aligned}$$

Combining, we have

$$\lambda(E) = \lim_n \sum_{k=1}^{n} \lambda(E_k) = \sum_{k=1}^{\infty} \lambda(E_k),$$

which proves that λ is countably additive. ■

Once we can construct another measure λ from a given measure μ , it is useful to think about their relationship. One useful relationship is that of *absolute continuity*.

Definition 10.4.1. Absolute Continuity Of A Measure

Let $(X, \mathcal{S}, \mu)$ be a measure space and let λ be another measure defined on $\mathcal{S}$. We say λ is absolutely continuous with respect to the measure μ if given E in $\mathcal{S}$ with $\mu(E) = 0$, then $\lambda(E) = 0$ also. This is written as $\lambda \ll \mu$.

We can also now prove an important result set within the framework of functions which are equal a.e.

Lemma 10.4.6. Function f Zero a.e. If and Only If Its Integral Is Zero

Let $(X, \mathcal{S}, \mu)$ be a measure space and let f be a function in $M^+(X, \mathcal{S})$. Then $f = 0$ a.e. if and only if $\int f d\mu = 0$.

Proof.
(⇐): If $\int f d\mu = 0$, then let $E_n = (f(x) > 1/n)$. Note $E_n \subseteq E_{n+1}$ so that (E_n) is an increasing sequence. Since (E_n) is an increasing sequence, we also know $\lim_n \mu(E_n) = \mu(\cup_n E_n)$. Further,

$$\cup_n E_n \;=\; \{x \,|\, f(x) > 0\}.$$

From the definition of E_n, we have

$$f(x) \;\geq\; \frac{1}{n}\, I_{E_n},$$

which implies

$$0 \;=\; \int f d\mu \;\geq\; \int \frac{1}{n}\, I_{E_n} \;=\; \frac{1}{n}\, \mu(E_n).$$

We see $\mu(E_n) \leq 1/n$ for all n which implies

$$\mu(f(x) > 0) \;=\; \lim_n \, \mu(E_n) \;\leq\; \lim_n \, 1/n \;=\; 0.$$

Hence, f is zero a.e.
(⇒): If f is zero a.e., let E be the set where $f(x) > 0$. Let $f_n = nI_E$. Note that

$$\liminf \, f_n(x) \;=\; \sup_m \, \inf_{n \geq m} \begin{cases} n & f(x) > 0 \\ 0 & f(x) = 0 \end{cases} \;=\; \sup_m \begin{cases} m & f(x) > 0 \\ 0 & f(x) = 0 \end{cases} \;=\; \begin{cases} \infty & f(x) > 0 \\ 0 & f(x) = 0. \end{cases}$$

Clearly, $f(x) \leq \liminf f_n(x)$ which implies $\int \, f \, d\mu \leq \int \, \liminf \, f_n \, d\mu$. Finally, by Fatou's Lemma, we find

$$\inf \, f \, d\mu \;\;\leq\;\; \int \liminf \, f_n \, d\mu \;\leq\; \liminf \int f_n \, d\mu \;=\; \liminf \, n\mu(E) \;=\; 0.$$

We conclude $\inf \, f \, d\mu = 0$. ■

Comment 10.4.1. *Given f in $M^+(X, \mathcal{S})$, Theorem 10.4.5 allows us to construct the new measure λ by $\lambda(E) = \int_E f d\mu$. If E has μ measure 0, we can use Lemma 10.4.6 to conclude that $\lambda(E) = 0$. Hence, a measure constructed in this way is absolutely continuous with respect to μ.*

We can now extend the Monotone Convergence Theorem slightly. It is often difficult to know that we have pointwise convergence up to a limit function on all of X. The next theorem allows us to relax the assumption to almost everywhere convergence as long as the underlying measure is complete.

Theorem 10.4.7. The Extended Monotone Convergence Theorem

Let $(X, \mathcal{S}, \mu)$ be a measure space with complete measure μ and let (f_n) be an increasing sequence of functions in $M^+(X, \mathcal{S})$. Let $f : X \rightarrow \overline{\Re}$ be an extended real valued function such that $f_n \rightarrow f$ pointwise a.e. on X. Then f is also in $M^+(X, \mathcal{S})$ and

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

Proof. *Let E be the set of points where f_n does not converge to f. Then by assumption E has measure 0 and $f_n \uparrow f$ on E^C. Thus,*

$$f_n \, I_{E^C} \;\uparrow\; f \, I_{E^C}$$

and applying the Monotone Convergence Theorem, we have

$$\lim_n \int f_n \, I_{E^C} = \int f \, I_{E^C}$$

and we can say $f \, I_{E^C}$ *is in* $M^+(X, \mathcal{S})$. *Now* f *is equal to* $f I_{E^C}$ *a.e. and so although in general,* f *need not be measurable, since* μ *is a complete measure, we can invoke Theorem 10.3.1 to conclude that* f *is actually measurable. Hence,* $f I_E$ *is measurable too. Since* $\mu(E) = 0$*, we thus know that*

$$\int f I_E \, d\mu = \int f_n I_E \, d\mu = 0.$$

Therefore, we have

$$\begin{aligned} \int f \, d\mu &= \int_E f \, d\mu + \int_{E^C} f \, d\mu = \int_{E^C} f \, d\mu \\ &= \lim_n \int_{E^C} f_n \, d\mu = \lim_n \left(\int_{E^C} f_n \, d\mu + \int_E f_n \, d\mu \right) = \lim_n \int f_n \, d\mu. \end{aligned}$$

■

10.5 Extending Integration To Extended Real Valued Functions

The results of the previous sections can now be used to extend the notion of integration to general extended real valued functions f in $M(X, \mathcal{S})$.

Definition 10.5.1. Summable Functions

Let $(X, \mathcal{S}, \mu)$ *be a measure space and* f *be in* $M(X, \mathcal{S})$. *We say* f *is summable or integrable on* X *if* $\int f^+ d\mu$ *and* $\int f^- d\mu$ *are both finite. In this case, we define the integral of* f *on* X *with respect to the measure* μ *to be*

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

Also, if E *is a measurable set, we define*

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

We let $L_1(X, \mathcal{S}, \mu)$ *be the collection of summable functions on* X *with respect to the measure* μ.

Comment 10.5.1. *If* f *can be decomposed into two non negative measurable functions* f_1 *and* f_2 *as* $f = f_1 - f_2$ *a.e. with* $\int f_1 d\mu$ *and* $\int f_2 d\mu$ *both finite, then note since* $f = f^+ - f^-$ *also, we have*

$$f_1 + f^- = f_2 + f^+.$$

Thus, since all functions involved are summable,

$$\int f_1 d\mu + \int f^- d\mu = \int f_2 d\mu + \int f^+ d\mu.$$

This implies that

$$\int (f_2 - f_1) d\mu \;=\; \int (f^+ - f^-) d\mu \;=\; \int f d\mu.$$

Hence, the value of the integral of f is independent of the decomposition.

There are a number of results that follow right away from this definition.

Theorem 10.5.1. Summable Implies Finite a.e.

Let $(X, \mathcal{S}, \mu)$ be a measure space and f be in $L_1(X, \mathcal{S})$. Then the set of points where f is not finite has measure 0.

Proof. *Let $E_n = (f(x) > n)$. Then it is easy to see that (E_n) is a decreasing sequence of sets and so*

$$\mu\Big(\bigcap_n E_n\Big) \;=\; \lim_n \mu(E_n).$$

It is also clear that

$$(f(x) = \infty) \;=\; \bigcap_n E_n.$$

Next, note

$$\begin{aligned}\int f^+ \, d\mu \;&=\; \int_{E_n} f^+ \, d\mu \;+\; \int_{E_n^C} f^+ \, d\mu \\ &\geq\; \int_{E_n} f^+ \, d\mu \;>\; n\, \mu(E_n).\end{aligned}$$

Thus, $\mu(E_n) < (\int f^+ \, d\mu)/n$. Since, the integral is a finite number, this tells us that $\lim_n \mu(E_n) = 0$. This immediately implies that $\mu(E) = 0$.

A similar argument shows that the set $(f(x) = -\infty)$ which is the same as the set $(f^-(x) = \infty)$ has measure 0. ■

Theorem 10.5.2. Summable Function Equal a.e. To Another Measurable Function Implies The Other Function Is Also Summable

Let $(X, \mathcal{S}, \mu)$ be a measure space and f be in $L_1(X, \mathcal{S})$. Then if $g \in M(X, \mathcal{S})$ with $f = g$ a.e., g is also summable.

Proof. *Let E be the set of points in X where f and g are not equal. Then E has measure zero. We then have $f I_{E^C} = g I_{E^C}$ and so $g I_{E^C}$ must be summable. Further, $f^+ I_{E^C} = g^+ I_{E^C}$ and $f^- I_{E^C} = g^- I_{E^C}$. We then note that*

$$\int g^+ I_{E^C} d\mu \;=\; \int g^+ I_{E^C} d\mu \;+\; \int g^+ I_E d\mu$$

because $\int g^+ I_E d\mu = 0$ since E has measure zero. But then we see

$$\int g^+ \, d\mu \;=\; \int g^+ I_{E^C} d\mu \;+\; \int g^+ I_E d\mu \;=\; \int f^+ I_{E^C} d\mu \;+\; \int f^+ I_E d\mu \;=\; \int f^+ \, d\mu.$$

Thus, we can see that $\int g^+ \, d\mu$ is finite. A similar argument shows $\int g^+ \, d\mu$ is finite and so g is summable. ■

Theorem 10.5.3. Summable Function Equal a.e. To Another Function With Measure Complete Implies The Other Function Is Also Summable

Let $(X, \mathcal{S}, \mu)$ be a measure space with μ complete and f be in $L_1(X, \mathcal{S})$. Then if g is a function equal a.e. to $f = g$, g is also summable.

Proof. *First, the completeness of μ implies that g is measurable. The argument to show g is summable is then the same as in the previous theorem's proof.* ■

We can extend the Monotone Convergence a bit more and actually construct a summable limit function in certain instances. This is known as Levi's Theorem.

Theorem 10.5.4. Levi's Theorem

Let $(X, \mathcal{S}, \mu)$ be a measure space and let (f_n) be a sequence of functions in $L_1(X, \mathcal{S}, \mu)$ which satisfy $f_n \leq f_{n+1}$ a.e. Further, assume

$$\lim_n \int f_n \, d\mu < \infty.$$

Then, there is a summable function f on X so that $f_n \uparrow f$ a.e. and $\int f_n d\mu \uparrow \int f d\mu$.

Proof. *Define the new sequence of functions (g_n) by $g_n = f_n - f_1$. Then, since (f_n) is increasing a.e., (g_n) is increasing and non negative a.e. By assumption, $\lim_n \int g_n d\mu$ is then finite. Call its value I for convenience of exposition. Now define the function g pointwise on X by*

$$g(x) \;=\; \lim_n g_n(x).$$

This limit always exists as an extended real number in $[0, \infty]$ and since each g_n is measurable, so is g. Let $E = (g(x) = \infty)$. Note that

$$E = \bigcap_i \left(\bigcup_n \Big(g_n(x) > i \Big) \right),$$

and so we know that E is measurable.

For each nonnegative measurable function g_i, there is an increasing sequence of simple functions (ϕ_n^i) such that $\phi_n^i \uparrow g_i$. For each n, define (recall the binary operator $\vee$ means a pointwise maximum)

$$\Psi_n \;=\; \phi_n^1 \,\vee\, \phi_n^2 \,\vee\, \cdots \,\vee\, \phi_n^n.$$

Then it is clear that Ψ_n is measurable. Given any x in X, we have that

$$\begin{aligned}
\Psi_{n+1}(x) &= \phi_{n+1}^1 \vee \phi_{n+1}^2 \vee \cdots \vee \phi_{n+1}^{n+1} \\
&\geq \phi_{n+1}^1 \vee \phi_{n+1}^2 \vee \cdots \vee \phi_n^{n+1} \\
&\geq \phi_n^1 \vee \phi_n^2 \vee \cdots \vee \phi_n^n \\
&= \Psi_n(x).
\end{aligned}$$

Hence, (Ψ_n) is an increasing sequence. Moreover, it is straightforward to see that

$$\Psi_n(x) \;\;\leq\;\; g_1(x) \vee g_2(x) \vee \cdots \vee g_n(x) \;\leq\; g_n(x) \;=\; g(x).$$

Hence, we know that $\lim_n \Psi_n(x) \leq g(x)$*. If this limit was strictly less than* $g(x)$*, let* r *denote half of the gap size; i.e.,* $r = (1/2)(g(x) - \lim_n \Psi_n(x)$*. Then, since* $\Psi_n(x) \geq \phi_n^i$ *where* i *is an index between* 1 *and* n*, we would have*

$$\phi_n^i \; < \; g(x) \; - \; r, \, 1 \; \leq \; i \; \leq \; n.$$

This implies that $\phi_n^n \leq g(x) - r$ *for all* n*. In particular, fixing the index* i*, we see that* $\phi_n^i \leq g(x) - r$ *for all* n*. But since* $\phi_n^i \uparrow g_i$*, this says* $g_i(x) \leq g(x) - r$*. Since, we can do this for all indices* i*, we have* $\lim_i g_i(x) \leq g(x) - r$ *or* $g(x) \leq g(x) - r$ *which is not possible. We conclude* $\lim_n \Psi_n = g$ *pointwise on* X*.*

Next, we claim $\int \Psi_n d\mu = \lim_n \int g_n d\mu$*. To see this, first notice that* $\int \Psi_n d\mu \geq \int \phi_n^i d\mu$ *for all* $1 \leq i \leq n$*. In fact, for any index* j*, there is an index* n^* *so that* $n^* > j$*. Hence,* $\int \Psi_{n^*} d\mu \geq \int \phi_n^j d\mu$*. This still holds for any* $n > n^*$ *as well. Thus, for any index* j*, we can say*

$$\lim_n \int \Psi_n \, d\mu \; \geq \; \lim_n \int \phi_n^j \, d\mu \; = \; \int g_j \, d\mu.$$

This implies that

$$\lim_n \int \Psi_n \, d\mu \; \geq \; \sup_j \int g_j \, d\mu \; = \; \lim_j \int g_j \, d\mu \; = \; I.$$

Also, since $\Psi_n \leq g_n(x)$*,*

$$\lim_n \int \Psi_n d\mu \; \leq \; \lim_n \int g_n d\mu \; = \; I.$$

This completes the proof that $\int \Psi_n d\mu = \lim_n \int g_n d\mu$*.*

We now show the measure of E *is zero. To do that, we start with the functions* $\Psi_n \wedge k I_E$ *for any positive integer* k*, where the wedge operation* $\wedge$ *is simply taking the minimum. If* $g(x)$ *is finite, then* $I_E(x) = 0$ *and since* Ψ_n *is non negative,* $\Psi_n \wedge k I_E = 0$*. On the other hand, if* $g(x) = \infty$*, then* $x \in E$ *and so* $k I_E(x) = k$*. Since* $\Psi_n \uparrow g$*, eventually,* $\Psi_n(x)$ *will exceed* k *and we will have* $\Psi_n \wedge k I_E = k$*. These two cases allow us to conclude*

$$\Psi_n \; \wedge \; k I_E \; \uparrow \; k I_E$$

for all x*. Thus,*

$$\int k I_E \, d\mu \; = \; \int \Psi_n \; \wedge \; k I_E \, d\mu \; \leq \; \int \Psi_n \, d\mu \; \leq \; \lim_n \int g_n d\mu \; = \; I.$$

We conclude $k \, \mu(E) \; \leq \; I$ *for all* k *which implies that* $\mu(E) = 0$*.*

Finally, to construct the summable function f *we need, define* $h \; = \; g I_{E^C}$*. Clearly,* $g_n \uparrow h$ *on* E^C*, that is, a.e. Also, since* $\Psi_n \uparrow g$ *on* E^C*, the Monotone Convergence Theorem tells us that*

$$\lim_n \int_{E^C} \Psi_n \, d\mu \; \uparrow \; \int_{E^C} g \, d\mu.$$

But,

$$\int_{E^C} g \, d\mu \; = \; \int h \, d\mu.$$

Hence, h is summable and so $f_1 + h$ is also summable. Define $f = f_1 + h$ on X and we have f is summable and

$$
\begin{aligned}
f_n \quad &\uparrow \quad f_1 + h \\
\int f_n \, d\mu \quad &= \quad \int f_1 \, d\mu + \int h \, d\mu \\
&= \quad \int f \, d\mu.
\end{aligned}
$$

■

Each summable function can also be used to construct a charge.

Theorem 10.5.5. Integrals Of Summable Functions Create Charges
Let $(X, \mathcal{S}, \mu)$ be a measure space and let f be a functions in $L_1(X, \mathcal{S}, \mu)$. Then the mapping $\lambda : \mathcal{S} \to \Re$ defined by

$$\lambda(E) \quad = \quad \int_E f \, d\mu$$

for all E in $\mathcal{S}$ defines a charge on $\mathcal{S}$. The integral $\int_E f \, d\mu$ is also called the indefinite integral of f with respect to the measure μ.

Proof. *Since f is summable, note that the mappings λ^+ and λ^- defined by*

$$\lambda^+(E) \quad = \quad \int_E f^+ \, d\mu, \ \ \lambda^-(E) \ = \ \int_E f^- \, d\mu$$

both define measures. It then follows immediately that λ is countably additive and hence is a charge. ■

Comment 10.5.2. *Since $\int_E f \, d\mu$ defines a charge and is countably additive, we see that if (E_n) is a collection of mutually disjoint measurable subsets, then*

$$\int_{\cup_n E_n} f \, d\mu \quad = \quad \sum_n \int_{E_n} f \, d\mu.$$

10.6 Properties Of Summable Functions

We need to know if $L_1(X, \mathcal{S}, \mu)$ is a linear space under the right interpretation of scalar multiplication and addition. To do this, we need some fundamental inequalities and conditions that force summability.

Theorem 10.6.1. Fundamental Abstract Integration Inequalities
Let $(X, \mathcal{S}, \mu)$ be a measure space.

(i): $f \in L_1(X, \mathcal{S}, \mu)$ if and only if $\mid f \mid \in L_1(X, \mathcal{S}, \mu)$.

(ii): $f \in L_1(X, \mathcal{S}, \mu)$ implies $\mid \int f \, d\mu \mid \leq \int \mid f \mid \, d\mu$.

(iii): f measurable and $g \in L_1(X, \mathcal{S}, \mu)$ with $\mid f \mid \leq \mid g \mid$ implies f is also summable and $\int \mid f \mid \, d\mu \leq \int \mid g \mid \, d\mu$.

Proof.
(i): If f is summable, f^+ and f^- are in $M^+(X, \mathcal{S}, \mu)$ with finite integrals. Since $\mid f \mid = f^+ + f^-$, we see $\mid f \mid^+ = \mid f \mid$ and $\mid f \mid^- = 0$. Thus, $\int \mid f \mid^+ d\mu = \int (f^+ + f^-) d\mu$ which is finite. Also, since $\int \mid f \mid^- d\mu = 0$, we see that $\mid f \mid$ is summable.

Conversely, if $\mid f \mid$ is summable, then $\int \mid f \mid^+ d\mu = \int (f^+ + f^-) d\mu$ is finite. This, in turn, tells us each piece is finite and hence f is summable too.
(ii): If f is summable, then

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

(iii): Since g is summable, so it $\mid g \mid$ by (i). Also, because $\mid f \mid \leq \mid g \mid$, each function is in $M^+(X, \mathcal{S})$ and so $\int \mid f \mid^+ d\mu \leq \int \mid g \mid^+ d\mu$ which is finite. Hence, $\mid f \mid$ is summable. Then, also by (i), f is summable. ■

We can now tackle the question of the linear structure of $L_1(X, \mathcal{S}, \mu)$.

Theorem 10.6.2. The Summable Function Form A Linear Space
Let $(X, \mathcal{S}, \mu)$ be a measure space. We define operations on $L_1(X, \mathcal{S}, \mu)$ as follows:

- *scalar multiplication: for all α in $\Re$, αf is the function defined pointwise by $(\alpha f)(x) = \alpha f(x)$.*
- *addition of functions: for any two functions f and g the sum of f and g is the new function defined pointwise on E_{fg}^C by $(f+g)(x) = f(x) + g(x)$, where, recall,*

$$E_{fg} = \left((f = \infty) \cap (g = -\infty) \bigcup (f = -\infty) \cap (g = \infty) \right).$$

This is equivalent to defining $f+g$ to be the function h where

$$h = (f+g) I_{E_{fg}^C}$$

This is a measurable function as we discussed in the proof of Lemma 9.6.3.

Then, we have

(i): αf is summable for all real α if f is summable and $\int \alpha f d\mu = \alpha \int f d\mu$.

(ii): $f+g$ is summable for all f and g which are summable and $\int (f+g) d\mu = \int f d\mu + \int g d\mu$.

Hence, $L_1(X, \mathcal{S}, \mu)$ is a vector space over $\Re$.

Proof.
(i): If α is 0, this is easy. Next, assume $\alpha > 0$. Then, $(\alpha f)^+ = \alpha f^+$ and $(\alpha f)^- = \alpha f^-$ and these two

functions are clearly summable since f^+ *and* f^- *are. Thus,* αf *is summable. Then, we have*

$$\begin{aligned}\int \alpha f \, d\mu &= \int (\alpha f)^+ \, d\mu - \int (\alpha f)^- \, d\mu \\ &= \alpha \Big(\int f^+ \, d\mu - \int f^- \, d\mu \\ &= \alpha \int f \, d\mu.\end{aligned}$$

Finally, if $\alpha < 0$*, we have* $(\alpha f)^+ = -\alpha f^-$ *and* $(\alpha f)^- = -\alpha f^+$*. Now simply repeat the previous arguments making a few obvious changes.*
(ii): Since f *and* g *are summable, we know that* $\mu(E_{fg} = 0$*. Further, we know* $\mid f \mid$ *and* $\mid g \mid$ *are summable. Since*

$$\mid f + g \mid \, I_{E^C_{fg}} \leq \Big(\mid f \mid + \mid g \mid \Big) \, I_{E^C_{fg}} \leq \mid f \mid + \mid g \mid,$$

we see $\mid f + g \mid \, I_{E^C_{fg}}$ *is summable by Theorem 10.6.1, part (iii). Hence,* $(f+g) \, I_{E^C_{fg}}$ *is summable also. Now decompose* $f+g$ *on* E^C_{fg} *as*

$$f + g = (f^+ + g^+) - (f^- + g^-).$$

Then, note

$$\begin{aligned}\int_{E^C_{fg}} (f+g) \, d\mu &= \int_{E^C_{fg}} (f^+ + g^+) \, d\mu - \int (f^- + g^-) \, d\mu \\ &= \int_{E^C_{fg}} (f^+ - f^-) \, d\mu + \int (g^+ - g^-) \, d\mu,\end{aligned}$$

where we are permitted to manipulate the terms in the integrals above because all are finite in value. However, we can rewrite this as

$$\int_{E^C_{fg}} (f+g) \, d\mu = \int_{E^C_{fg}} f \, d\mu + \int g \, d\mu.$$

Since we define the sum of f *and* g *to be the function* $(f+g) \, I_{E^C_{fg}}$*, we see* $f+g$ *is in* $L_1(X, \mathcal{S}, \mu)$*.* ∎

10.7 The Dominated Convergence Theorem

We can now complete this chapter by proving the important limit interchange called the *Lebesgue Dominated Convergence Theorem.*

Theorem 10.7.1. Lebesgue's Dominated Convergence Theorem

Let $(X, \mathcal{S}, \mu)$ *be a measure space,* (f_n) *be a sequence of functions in* $L_1(X, \mathcal{S}, \mu)$ *and* $f : X \to \overline{\Re}$ *so that* $f_n \to f$ *a.e. Further, assume there is a summable* g *so that* $\mid f_n \mid \leq g$ *for all* n*. Then, suitably defined,* f *is also measurable and summable with* $\lim_n \int f_n \, d\mu = \int f \, d\mu$*.*

Proof. *Let E be the set of points in X where the sequence does not converge. Then, by assumption, $\mu(E) = 0$ and*

$$f_n I_{E^C} \quad \rightarrow \quad f I_{E^C}, \quad \textit{and} \quad | \, f_n I_{E^C} \, | \leq \, g I_{E^C}.$$

Hence, $| \, f I_{E^C}$ is measurable and satisfies $| \, f I_{E^C} \, | \leq g I_{E^C}$. Therefore, since g is summable, we have that $f I_{E^C}$ is summable too.

We can write out our fundamental inequality as follows

$$- g I_{E^C} \; \leq \; f_n \, I_{E^C} \quad \leq \quad g I_{E^C}. \tag{$\boldsymbol{\alpha}$}$$

This implies that $h_n = f_n \, I_{E^C} \, + \, g I_{E^C}$ is non negative and hence, we can apply Fatou's lemma to find

$$\int \liminf \, h_n \, d\mu \quad \leq \quad \liminf \int h_n \, d\mu.$$

However, we know

$$\begin{aligned} \liminf \, h_n \quad &= \quad \liminf \Big(f_n \, I_{E^C} \, + \, g I_{E^C} \Big) \\ &= \quad g I_{E^C} \, + \, \liminf \, f_n \, I_{E^C} \\ &= \quad g I_{E^C} \, + \, f \, I_{E^C}, \end{aligned}$$

because f_n converges pointwise to f on E^C. It then follows that

$$\begin{aligned} \int \Big(g I_{E^C} \, + \, f \, I_{E^C} \Big) \, d\mu \quad &\leq \quad \liminf \int \Big(f_n \, I_{E^C} \, + \, g I_{E^C} \Big) \, d\mu \\ &= \quad \int g I_{E^C} \, d\mu \, + \, \liminf \int f_n \, I_{E^C} \, d\mu. \end{aligned}$$

Since g is summable, we also know

$$\int (g \, I_{E^C} \, + \, f \, I_{E^C}) \, d\mu \quad = \quad \int (g \, I_{E^C} \, d\mu \, + \int f \, I_{E^C}) \, d\mu.$$

Using this identity, we have

$$\int (g \, I_{E^C} \, d\mu \, + \int f \, I_{E^C}) \, d\mu. \quad \leq \quad \int g \, I_{E^C} \, d\mu \, + \, \liminf \int f_n \, I_{E^C} \, d\mu.$$

The finiteness of the integral of the g term then allows cancellation so that we obtain the inequality

$$\int f \, I_{E^C} \, d\mu. \quad \leq \quad \liminf \int f_n \, I_{E^C} \, d\mu.$$

Since the integrals of f and f_n are all zero on E, we have shown

$$\int f \, d\mu. \quad \leq \quad \liminf \int f_n \, d\mu.$$

We now show the reverse inequality holds. Using Equation $\boldsymbol{\alpha}$*, we see* $z_n = gI_{E^C} - f_n I_{E^C}$ *is also non negative for all* n*. Applying Fatou's Lemma, we find*

$$\int \liminf \, z_n \, d\mu \;\; \leq \;\; \liminf \int \, z_n \, d\mu.$$

Then, we note

$$\begin{aligned} \liminf \, z_n &= \liminf \Big(-f_n \, I_{E^C} \, + \, gI_{E^C} \Big) \\ &= gI_{E^C} \, + \, \liminf \Big(-f_n \, I_{E^C} \Big) \\ &= gI_{E^C} \, - \, f \, I_{E^C}, \end{aligned}$$

because f_n *converges pointwise to* f *on* E^C*. It then follows that*

$$\begin{aligned} \int \Big(gI_{E^C} \, - \, f \, I_{E^C} \Big) \, d\mu &\leq \liminf \int \Big(-f_n \, I_{E^C} \, + \, gI_{E^C} \Big) \, d\mu \\ &= \int \, gI_{E^C} \, d\mu \, + \, \liminf \int \Big(-f_n \, I_{E^C} \Big) \, d\mu. \end{aligned}$$

Now,

$$\begin{aligned} \liminf \int \Big(-f_n \, I_{E^C} \Big) \, d\mu &= \sup_m \inf_{m \geq n} \int \Big(-f_n \, I_{E^C} \Big) \, d\mu \\ &= \sup_m \Big(- \sum_{m \geq n} \int \, f_n \, I_{E^C} \, d\mu \Big) \\ &= - \inf_m \sup_{m \geq n} \int \, f_n \, I_{E^C} \, d\mu \\ &= - \limsup \int \, f_n \, I_{E^C} \, d\mu. \end{aligned}$$

Thus, we can conclude

$$\int \Big(gI_{E^C} \, - \, f \, I_{E^C} \Big) \, d\mu \;\; \leq \;\; \int \, gI_{E^C} \, d\mu \, - \, \limsup \int \, f_n \, I_{E^C} \, d\mu.$$

Again, since g *is summable, we can write*

$$\int \, gI_{E^C} \, d\mu - \int \, f \, I_{E^C} \, d\mu \;\; \leq \;\; \int \, gI_{E^C} \, d\mu \, - \, \limsup \int \, f_n \, I_{E^C} \, d\mu.$$

After canceling the finite value $\int gI_{E^C} \, d\mu$*, we have*

$$\int \, f \, I_{E^C} \, d\mu \;\; \geq \;\; \limsup \int \, f_n \, I_{E^C} \, d\mu.$$

This then implies, using arguments similar to the ones used in the first case, that

$$\int \, f \, d\mu \;\; \geq \;\; \limsup \int \, f_n \, d\mu.$$

However, limit inferiors are always less than limit superiors and so we have

$$\limsup \int f_n \, d\mu \quad \leq \quad \int f \, d\mu \; \leq \; \liminf \int f_n \, d\mu \; \leq \; \limsup \int f_n \, d\mu.$$

It follows immediately that $\lim_n \int f_n d\mu = \int f d\mu$.

Finally, we can now see how to define f in a suitable fashion. The function $f I_{E^C}$ is measurable and is 0 on E. Hence, the limit function f can has the form

$$f(x) \quad = \quad \begin{cases} \lim_n f_n(x) & \text{when the limit exists, i.e. when } x \in E^C \\ 0 & \text{when the limit does not exist, i.e. when } x \in E. \end{cases}$$

■

10.8 Homework

Exercise 10.8.1. *Assume $f \in L_1(X, \mathcal{S}, \mu)$ with $f(x) > 0$ on X. Further, assume there is a positive number α so that $\alpha < \mu(X) < \infty$. Prove that*

$$\inf \{ \int_E f \, d\mu \mid \mu(E) \; \geq \alpha \} \; > \; 0.$$

Exercise 10.8.2. *Assume $f \in L_1(X, \mathcal{S}, \mu)$. Let $\alpha > 0$. Prove that*

$$\mu(\{x \mid \; | \, f(x) \, | \geq \alpha\})$$

is finite.

Exercise 10.8.3. *Assume $(f_n) \subseteq L_1(X, \mathcal{S}, \mu)$. Let $f : X \to \overline{\Re}$ be a function. Assume $f_n \to f[ptws\ ae]$. Prove*

$$\int \mid f_n - f \mid \; d\mu \; \to \; 0 \Rightarrow \int \mid f_n \mid \; d\mu \; \to \; \int \mid f \mid \; d\mu$$

Exercise 10.8.4. *Let $(X, \mathcal{S})$ be a measurable space. Let $\mathcal{C}$ be the collection of all charges on $\mathcal{S}$. Prove that $\mathcal{C}$ is a Banach Space under the operations*

$$\begin{aligned} \Big(c\,\mu\Big)(E) \quad &= \quad c\,\mu(E), \; \forall \, c \; \in \; \Re, \; \forall \, \mu \\ \Big(\mu \; + \; \nu\Big)(E) \quad &= \quad \mu(E) \; + \; \nu(E), \; \forall \, \mu, \; \nu, \end{aligned}$$

with norm $\| \; \mu \; \| = \; |\mu|(X)$

Chapter 11

The $\mathcal{L}_p$ Spaces

In mathematics and other fields, we often group objects of interest into sets and study the properties of these sets. In this book, we have been studying a set X with a sigma - algebra of subsets contained within it, the collection of functions which are measurable with respect to the sigma - algebra and recently, the set of functions which are summable. In addition, we have noted that the sets of measurable and summable functions are closed under scalar multiplication and addition as long as we interpret addition in the right way when the functions are extended real - valued.

We can do more along these lines. We will now study the sets of summable functions as vector spaces with a suitable norm. We begin with a review.

Definition 11.0.1. The Norm On A Vector Space

Let X be a non empty vector space over $\Re$. We say $\rho : X \rightarrow \Re$ is a norm on X if

(N1): $\rho(x)$ is non negative for all x in X,

(N2): $\rho(x) = 0 \Leftrightarrow x = 0$,

(N3): $\rho(\alpha x) = |\alpha| \rho(x)$, for all α in $\Re$ and for all x in X,

(N4): $\rho(x + y) \leq \rho(x) + \rho(y)$, for all x and y in X.

If ρ satisfies only N1, N3 and N4, we say ρ is a semi-norm or pseudo-norm. We will usually denote a norm of x by the symbol $\| x \|$.
The pair $(X, \| \|)$ is called a Normed Linear Space or NLS.

If a set X has no linear structure, we can still have a notion of the distance between objects in the set, if the set is endowed with a metric. This is defined below.

Definition 11.0.2. The Metric On A Set

Let X be a non empty set. We say $d : X \times X \rightarrow \Re$ is a metric if

(M1): $d(x, y)$ is non negative for all x and y in X,

(M2): $d(x, y) = 0 \Leftrightarrow x = y$,

(M3): $d(x, y) = d(y, x)$, for all for all x and y in X,

(M4): $d(x, y) \leq d(x, z) + d(y, z)$, for all x, y and z in X.

If d satisfies only M1, M2 and M4, we say d is a semi-metric or pseudo-metric. The pair (X, d) is called a metric space. Note that in a metric space, there is no notion of scaling or adding objects because there is no linear structure.

Comment 11.0.1. *It is a standard result from a linear analysis course, that the norm in a NLS $(X, \| \|)$ induces a metric on X by defining*

$$d(x, y) = \| x - y \|, \ \forall\, x,\ y \in X.$$

Given a sequence (x_n) in a NLS $(X, \| \|)$, we can define what we mean by the convergence of this sequence to another object x in X.

Definition 11.0.3. Norm Convergence

Let $(X, \| \|)$ be a non empty NLS. Let (x_n) be a sequence in X. We say the sequence (x_n) converges to x in X if

$$\forall\, \epsilon > 0, \exists\, N \ \ni n > N \Rightarrow \| x_n - x \| < \epsilon.$$

Now let $(X, \mathcal{S}, \mu)$ be a nonempty measurable space. We are now ready to discuss the space $L_1(X, \mathcal{S}, \mu)$. By Theorem 10.6.2, we know that this space is a vector space with suitably defined addition. We can now define a semi-norm for this space.

Theorem 11.0.1. The L_1 Semi-norm

Let $(X, \mathcal{S}, \mu)$ be a nonempty measurable space. Define $\| x \|_1$ on $L_1(X, \mathcal{S}, \mu)$ by

$$\| f \|_1 \ \int |f|\, d\mu, \ \forall\, f \in L_1(X, \mathcal{S}, \mu).$$

Then, $\| x \|_1$ is a semi-norm. Moreover, property N3 is almost satisfied: instead of N3, we have

$$\| f \|_1 = 0 \Leftrightarrow f = a.e.$$

Proof.

(N1): $\| f \|_1$ is clearly non negative.

(N2): This proof is an easy calculation.

$$\begin{aligned} \| \alpha f \|_1 &= \int |\alpha\, f|\, d\mu = \int |\alpha|\, |f|\, d\mu \\ &= |\alpha| \int |f|\, d\mu = |\alpha| \ \| f \|_1 . \end{aligned}$$

(N4): To prove this, we start with the triangle inequality for real numbers. We know that if f and g are summable, then the sum of $f+g$ is defined to be $h = (f+g)I_{E_{fg}^C}$. Let A be the set of points where this sum is ∞ and B be the set where the sum if $-\infty$. Then on $\mu(E_{fg} \cup A \cup B) = 0$ and on $(E_{fg} \cup A \cup B)^C$, h is finite. For convenience of exposition, we will simply write h as $f+g$ from now on. So $f+g$ is finite off a set of measure 0. At the points where $f+g$ is finite, we can apply the standard triangle inequality to $f(x)+g(x)$. We have

$$|f(x)+g(x)| \le |f(x)| + |g(x)|, a.e.$$

This implies

$$\int |f+g|\, d\mu \le \int |f|\, d\mu + \int |g|\, d\mu.$$

At the risk of repeating ourselves too much, let's go through the integral on the left hand side again. We actually have

$$\begin{aligned} \int |f+g|\, I_{E_{fg}^C \cap A^C \cap B^C}\, d\mu &= \int h\, I_{A^C \cap B^C}\, d\mu \\ &= \int h\, d\mu \end{aligned}$$

since $\mu(A^C \cap B^C) = 0$. Now the above inequality estimates clearly tell us

$$\| f+g \|_1 \le \| f \|_1 + \| g \|_1 .$$

Finally, we look at what is happening in condition N2. Since $|f|$ is in $M^+(X, \mathcal{S}, \mu)$, by Lemma 10.4.6, we know

$$|f| = 0 a.e. \Leftrightarrow \int |f|\, d\mu = 0.$$

Hence, $\| f \|_1 = 0$ if and only if $f = 0$ a.e. ■

Although $\| x \|_1$ is only a semi-norm, there is a way to think of this class of functions as a normed linear space. Let's define two functions f and g in $L_1(X, \mathcal{S}, \mu)$ to be equivalent or to be precise, μ - equivalent if $f = g$ except of a set of μ measure 0. We use the notation $f \sim g$ to indicate this equivalence. It is easy to see that $\sim$ defines an equivalence relation on $L_1(X, \mathcal{S}, \mu)$. We will let $[\boldsymbol{f}]$ denote the equivalence class defined by f:

$$[\boldsymbol{f}] = \{g \in L_1(X, \mathcal{S}, \mu) \mid g \sim f\}.$$

Any g in $[\boldsymbol{f}]$ is called a representative of the equivalence class $[\boldsymbol{f}]$. A straightforward argument shows that two equivalence classes $[\boldsymbol{f_1}]$ and $[\boldsymbol{f_2}]$ are either equal as sets or disjoint. The collection of all distinct equivalence classes of $L_1(X, \mathcal{S}, \mu)$ under a.e. equivalence will be denoted by $\mathcal{L}_1(X, \mathcal{S}, \mu)$.

Theorem 11.0.2. $\mathcal{L}_1$ **Is A Normed Linear Space**
$\mathcal{L}_1(X, \mathcal{S}, \mu)$ is a vector space over $\Re$ with scalar multiplication and object addition defined as

$$\begin{aligned} \alpha\,[\boldsymbol{f}] &= [\boldsymbol{\alpha}\,\boldsymbol{f}]\,\forall\,[\boldsymbol{f}] \\ [\boldsymbol{f}] + [\boldsymbol{g}] &= [\boldsymbol{f} + \boldsymbol{g}], \forall[\boldsymbol{f}] \text{ and } [\boldsymbol{g}]. \end{aligned}$$

Further, $\|\,[\boldsymbol{f}]\,\|_1$ defined by

$$\|\,[\boldsymbol{f}]\,\|_1 = \int |g|\,d\mu,$$

for any representative g of $[\boldsymbol{f}]$ is a norm on $\mathcal{L}_1(X, \mathcal{S}, \mu)$.

Proof. *The definition of scalar multiplication is clear. However, as usual, we can spend some time with addition. We know $f + g$ is defined on E_{fg}^C and that E_{fg} has measure 0. Hence, if $u \in [\boldsymbol{f}]$ and $v \in eclassg$, then $u = f$ and $v = g$ except on sets A and B of measure 0. Also, as usual, the sum $u + v$ is defined on E_{uv}^C. Hence,*

$$u + v = f + g, \; x \in E_{uv}^C \cap E_{fg}^C \cap A^C \cap B^C.$$

which is the complement of a set of measure 0. Hence, $u + v \in [\boldsymbol{f} + \boldsymbol{g}]$. Thus, $[\boldsymbol{f}] + [\boldsymbol{f}] \subseteq [\boldsymbol{f} + \boldsymbol{g}]$. Conversely, let $h \in [\boldsymbol{f} + \boldsymbol{g}]$. Now

$$(f + g)\, I_{Efg^C} = f\, I_{Efg^C} + g\, I_{Efg^C}.$$

Hence, if we let

$$u = f\, I_{Efg^C} \text{ and } v = g\, I_{Efg^C},$$

we see $h \sim (u + v)$, with $u \in [\boldsymbol{f}]$ and $v \in [\boldsymbol{g}]$. We conclude $[\boldsymbol{f} + \boldsymbol{g}] \subseteq [\boldsymbol{f}] + [\boldsymbol{g}]$. Hence, the addition of equivalence classes makes sense.

We now turn our attention to the possible norm $\|\,[\boldsymbol{f}]\,\|_1$. First, we must show that our definition of norm is independent of the choice of representative chosen from $[\boldsymbol{f}]$. If $g \sim f$, then $g = f$ except on a set A of measure 0. Thus, we know the integral of f and g match by Lemma 10.4.6. Here are the details:

$$\begin{aligned} \int |g|\,d\mu &= \int_A |g|\,d\mu + \int_{A^C} |g|\,d\mu \\ &= 0 + \int_{A^C} |f|\,d\mu \\ &= \int_A |f|\,d\mu + \int_{A^C} |f|\,d\mu \\ &= \int |f|\,d\mu. \end{aligned}$$

We conclude the value of $\|\,[\boldsymbol{f}]\,\|_1$ is independent of the choice of representative from $[\boldsymbol{f}]$. Now we prove this is a norm.
(N1): $\|\,[\boldsymbol{f}]\,\|_1 = \int |g| d\mu \geq 0$.
(N2): If $\|\,[\boldsymbol{f}]\,\|_1 = 0$, then for any representative g of $[\boldsymbol{f}]$, we have $\int |g| d\mu = 0$. By Lemma 10.4.6, this implies that $g = 0$ a.e. and hence, $g \in [\boldsymbol{0}]$ (we abuse notation here by simply writing the zero function

$h(x) = 0$, $\forall x$ as 0). But since $g \in [\boldsymbol{f}]$ also, this means $[\boldsymbol{f}] \cap [\boldsymbol{0}]$ is nonempty. This immediately implies that $[\boldsymbol{f}] = [\boldsymbol{0}]$. Conversely, if $[\boldsymbol{f}] = [\boldsymbol{0}]$, the result is clear. (N3): Let α be a real number. Then, if g is any representative of $[\boldsymbol{f}]$, we have $\alpha\, g$ is a representative of $[\boldsymbol{\alpha}\ \boldsymbol{f}]$. We find

$$\begin{aligned} \| \, [\boldsymbol{\alpha}\ \boldsymbol{f}] \, \|_1 &= \int |\alpha\, g| \, d\mu \;=\; |\alpha| \int |g| \, d\mu \\ &= |\alpha| \;\; \| \, [\boldsymbol{f}] \, \| \, . \end{aligned}$$

(N4): The triangle inequality follows from the triangle inequality that holds for the representatives. ■

11.1 The General L_p spaces

We can construct additional spaces of summable functions. Let p be a real number satisfying $1 \leq p < \infty$. Then the function $\phi(u) = u^p$ is a continuous function on $[0, \infty)$ that satisfies $\lim_n \phi(n) = \infty$. Thus, by Lemma 9.6.9, if f is an extended real - valued function on X, then the composition $\phi \circ |f|$ or $|f|^p$ is also measurable. Hence, we know the integral $\int |f|^p \, d\mu$ exists as an extended real - valued number. The class of measurable functions that satisfy $\int |f|^p \, d\mu < \infty$ is another interesting class of functions.

We begin with some definitions.

Definition 11.1.1. The Space Of p Summable Functions

$(X, \mathcal{S}, \mu)$ be a nonempty measurable space. Let p be a real number satisfying $1 \leq p < \infty$. Then, $|f|^p$ is a measurable function. We let

$$L_p(X, \mathcal{S}, \mu) \;=\; \{\; f \in M(X, \mathcal{S}, \mu) \mid \int |f|^p \, d\mu < \infty.$$

For later use, we will also define what are called *conjugate index pairs.*

Definition 11.1.2. Conjugate Index Pairs

Let p be a real number satisfying $1 \leq p \leq \infty$. If $1 < p$ is finite, the index conjugate to p is the real number q satisfying

$$\frac{1}{p} + \frac{1}{q} = 1,$$

while if $p = 1$, the index conjugate to p is $q = \infty$.

We will be able to show that $L_p(X, \mathcal{S}, \mu)$ is a vector space under the usual scalar multiplication and addition operations once we prove some auxiliary results. These are the Hölder's and Minkowski's Inequality. First, there is a standard lemma we will call the *Real Number Conjugate Indices Inequality.*

Lemma 11.1.1. Real Number Conjugate Indices Inequality

Let $1 < p < \infty$ and q be the corresponding conjugate index. Then if α and β are positive numbers,

$$A\, B \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

Proof. *This proof is standard in any Linear Analysis book and so we will not repeat it here.* ■

Theorem 11.1.2. Hölder's Inequality

Let $1 < p < \infty$ and q be the index conjugate to p. Let $f \in L_p(X, \mathcal{S}, \mu)$ and $g \in L_q(X, \mathcal{S}, \mu)$. Then $f\,g \in L_1(X, \mathcal{S}, \mu)$ and

$$\int |fg|\, d\mu \leq \left(\int |f|^p\, d\mu\right)^{1/p} \left(\int |g|^q\, d\mu\right)^{1/q}.$$

Proof. *The result is clearly true if $f = g = 0$ a.e. Also, if $\int |f|^p d\mu = 0$, then $|f|^p = 0$ a.e. which tells us $f = 0$ a.e. and the result follows again. We handle the case where $\int |g|^q d\mu = 0$ in a similar fashion. Thus, we will assume both $I^p = \int |f|^p d\mu > 0$ and $J^q = \int |g|^q d\mu > 0$.*

Let E_f and E_g be the sets where f and g are not finite. By our assumption, we know the measure of these sets is 0. Hence, for all x in $E_f^C \cap E_g^C$, the values $f(x)$ and $g(x)$ are finite. We apply Lemma 11.1.1 to conclude

$$\frac{|f(x)|}{I}\,\frac{|g(x)|}{J} \;\leq\; (1/p)\,\frac{|f(x)|^p}{I^p} \;+\; (1/q)\,\frac{|g(x)|^q}{J^q}.$$

holds on $E_f^C \cap E_g^C$. Off of this set, we have that the left hand side is ∞ and so is the left hand side. Hence, even on $E_f \cup E_g$, the inequality is satisfied. Thus, since the function on the right hand side is summable, we must have the left hand side is a summable function too by Theorem 10.6.1. Hence, $f\,g \in L_1(X, \mathcal{S}, \mu)$. We then have

$$\begin{aligned}
\int \frac{|f(x)|}{I}\,\frac{|g(x)|}{J}\, d\mu \;&\leq\; \int (1/p)\,\frac{|f(x)|^p}{I^p}\, d\mu \;+\; \int (1/q)\,\frac{|g(x)|^q}{J^q}\, d\mu \\
&=\; \frac{1}{pI^p}\int |f(x)|^p\, d\mu \;+\; \frac{1}{qJ^q}\int |g(x)|^q\, d\mu \\
=\; \frac{1}{p} + \frac{1}{q} \;=\; 1.
\end{aligned}$$

Thus,

$$\int |f\,g|\, d\mu \;\leq\; I\,J \;=\; \left(\int |f|^p\, d\mu\right)^{1/p} \left(\int |g|^q\, d\mu\right)^{1/p}.$$

∎

The special case of $p = q = 2$ is of great interest. The resulting Hölder's Inequality is often called the Cauchy - Schwartz Inequality. We see

Theorem 11.1.3. Cauchy - Bunyakovskiĭ - Schwartz Inequality

Let $f, g \in L_2(X, \mathcal{S}, \mu)$. Then $f\,g \in L_1(X, \mathcal{S}, \mu)$ and

$$\int |fg|\, d\mu \leq \left(\int |f|^2\, d\mu\right)^{1/2} \left(\int |g|^2\, d\mu\right)^{1/2}.$$

Theorem 11.1.4. Minkowski's Inequality

Let $1 \leq p < \infty$ and let $f, g \in L_p(X, \mathcal{S}, \mu)$. Then $f+g$ is in $L_p(X, \mathcal{S}, \mu)$ and

$$\left(\int |f\,+\,g|^p\,d\mu\right)^{1/p} \leq \left(\int |f|^p\,d\mu\right)^{1/p} + \left(\int |g|^p\,d\mu\right)^{1/p}.$$

Proof. *If $p=1$, this is property N4 of the semi-norm $\|\cdot\|_1$. Thus, we can assume $1 < p < \infty$. Since f and g are measurable, we define the sum of $f+g$ as $h = (f+g)\,I_A$ where $A = E_{fg}^C$ with $\mu(E_{fg} = 0$. Then as discussed h is measurable. We see on A,*

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

even when function values are ∞. Hence,

$$|f(x)+g(x)|^p \;\leq\; 2^p\left(\max\{|f(x)|, |g(x)|\}\right)^p \;\leq\; 2^p\left(|f(x)|^p + |g(x)|^p\right).$$

Then, since the right hand side is summable, so is the left hand side. We conclude $f+g$ is in $L_p(X, \mathcal{S}, \mu)$. Note this also tells us $|f+g|$ is in $L_1(X, \mathcal{S}, \mu)$. Further,

$$|f(x)+g(x)|^p \quad = \quad |f+g|\,|f+g|^{p-1} \;\leq\; |f|\,|f+g|^{p-1} \,+\, |g|\,|f+g|^{p-1}.$$

We have the identity

$$|f(x)+g(x)|^p \quad \leq \quad |f|\,|f+g|^{p-1} \,+\, |g|\,|f+g|^{p-1}. \tag{$*$}$$

Now since p and q are conjugate indices, we know

$$\begin{aligned}(1/p) \,+\, (1/q) \,=\, 1 \quad &\Rightarrow \quad p \,+\, q \,=\, pq\\ &\Rightarrow \quad p \,=\, q(p-1).\end{aligned}$$

Thus, the function

$$\left(\left|f+g\right|^{p-1}\right)^q \quad = \quad |f+g|^p,$$

and so this function is summable implying $|f+g|^{p-1} \in L_q(X, \mathcal{S}, \mu)$. Now apply Hölder's Inequality to the two parts of the right hand side of Equation $*$. We find

$$\int |f|\,|f+g|^{p-1}\,d\mu \;\leq\; \left(\int |f|^p\,d\mu\right)^{1/p}\left(\int \left(|f+g|^{p-1}\right)^q d\mu\right)^{1/q}.$$

and

$$\int |g|\,|f+g|^{p-1}\,d\mu \;\leq\; \left(\int |g|^p\,d\mu\right)^{1/p}\left(\int \left(|f+g|^{p-1}\right)^q d\mu\right)^{1/q}.$$

But we have learned we can rewrite the second terms of the above inequalities to get

$$\int |f|\,|f+g|^{p-1}\,d\mu \le \left(\int |f|^p\,d\mu\right)^{1/p} \left(\int \left(|f+g|^p\,d\mu\right)^{1/q}\right).$$

and

$$\int |g|\,|f+g|^{p-1}\,d\mu \le \left(\int |g|^p\,d\mu\right)^{1/p} \left(\int \left(|f+g|^p\,d\mu\right)^{1/q}\right).$$

Thus, combining, we have

$$\int |f+g|^p\,d\mu \le \left(\left(\int |f|^p\,d\mu\right)^{1/p} + \left(\int |g|^p\,d\mu\right)^{1/p}\right) \left(\int \left(|f+g|^p\,d\mu\right)^{1/q}\right).$$

We can rewrite this as

$$\left(\int |f+g|^p\,d\mu\right)^{1-1/q} \le \left(\int |f|^p\,d\mu\right)^{1/p} + \left(\int |g|^p\,d\mu\right)^{1/p}.$$

Since $1-1/q = 1/p$, we have established the desired result. ∎

Hölder's and Minkowski's Inequalities allow us to prove that the L_p spaces are normed linear spaces.

Theorem 11.1.5. L_p Is A Vector Space

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $1 \le p < \infty$. Then, if scalar multiplication and object addition are defined pointwise as usual, $L_p(X, \mathcal{S}, \mu)$ is a vector space.

Proof. *The only thing we must check is that if f and g are in $L_p(X, \mathcal{S}, \mu)$, then so is $f+g$. This follows from Minkowski's inequality.* ∎

Since $L_p(X, \mathcal{S}, \mu)$ is a vector space, the next step is to find a norm for the space.

Theorem 11.1.6. The L_p Semi-Norm

Let $(X, \mathcal{S}, \mu)$ be a measure space and let $1 \le p < \infty$. Define $\|\cdot\|_p$ on $L_p(X, \mathcal{S}, \mu)$ by

$$\| f \|_p = \left(\int |f|^p\,d\mu\right)^{1/p}.$$

Then, $\|\cdot\|_p$ is a semi-norm.

Proof. *Properties N1 and N3 of a norm are straightforward to prove. To see that the triangle inequality holds, simply note that Minkowski's Inequality can be rewritten as*

$$\| f+g \|_p \le \| f \|_p + \| g \|_p,$$

for arbitrary f and g in $L_p(X, \mathcal{S}, \mu)$. ∎

If we use the same notion of equivalence a.e. as did earlier, we can define the the collection of all distinct equivalence classes of $L_p(X, \mathcal{S}, \mu)$ under a.e. equivalence. This will be denoted by $\mathcal{L}_p(X, \mathcal{S}, \mu)$. We can prove that this space is a normed linear space using the norm $\| [\cdot] \|_p$.

Theorem 11.1.7. $\mathcal{L}_p$ Is A Normed Linear Space

Let $1 \leq p < \infty$. Then $\mathcal{L}_p(X, \mathcal{S}, \mu)$ is a vector space over $\Re$ with scalar multiplication and object addition defined as

$$\begin{aligned} \alpha\,[\boldsymbol{f}] &= [\boldsymbol{\alpha}\,\boldsymbol{f}]\,\forall\,[\boldsymbol{f}] \\ [\boldsymbol{f}] + [\boldsymbol{g}] &= [\boldsymbol{f} + \boldsymbol{g}], \forall [\boldsymbol{f}] \text{ and } [\boldsymbol{g}]. \end{aligned}$$

Further, $\| [\boldsymbol{f}] \|_p$ defined by

$$\| [\boldsymbol{f}] \|_p = \left(\int |g|^p \, d\mu \right)^{1/p},$$

for any representative g of $[\boldsymbol{f}]$ is a norm on $\mathcal{L}_p(X, \mathcal{S}, \mu)$.

Proof. *The proof of this is quite similar to that of Theorem 11.0.2 and so we will not repeat most of it.* ■

We will now show that $\mathcal{L}_p(X, \mathcal{S}, \mu)$ is a complete NLS. First, recall what a Cauchy Sequence means.

Definition 11.1.3. Cauchy Sequence In Norm

Let $(X, \|\cdot\|)$ be a NLS. We say the sequence (f_n) of X is a Cauchy Sequence, if given $\epsilon > 0$, there is a positive integer N so that

$$\| f_n - f_m \| < \epsilon, \; \forall\, n, \, m > N.$$

This leads to the definition of a complete NLS or Banach space.

Definition 11.1.4. Complete NLS

Let $(X, \|\cdot\|)$ be a NLS. We say the X is a complete NLS if every Cauchy sequence in X converges to some object in X.

It is a standard proof to show that any sequence in a NLS that converges must be a Cauchy sequence. Let's prove that in the context of the $\mathcal{L}_p(X, \mathcal{S}, \mu)$ space to get some practice.

Theorem 11.1.8. Sequences That Converge in $\mathcal{L}_p$ Are Cauchy

Let $([\boldsymbol{f_n}])$ be a sequence in $\mathcal{L}_p(X, \mathcal{S}, \mu)$ which converges to $[\boldsymbol{f}]$ in $\mathcal{L}_p(X, \mathcal{S}, \mu)$ in the $\|\cdot\|_p$ norm. Then, $([\boldsymbol{f_n}])$ is a Cauchy sequence.

Proof. *Let $\epsilon > 0$ be given. Then, there is a positive integer N so that if $n > N$, then*

$$\| [\boldsymbol{f_n} - \boldsymbol{f}] \|_p < \epsilon/2.$$

Thus, if n and m are larger than N, we by property N4 that

$$\| [\boldsymbol{f_n} - \boldsymbol{f_m}] \|_p \quad = \quad \| [(\boldsymbol{f_n} - \boldsymbol{f}) \; + \; (\boldsymbol{f} - \boldsymbol{f_m})] \|_p \leq \| [\boldsymbol{f_n} - \boldsymbol{f}] \|_p \; + \; \| [\boldsymbol{f_m} - \boldsymbol{f}] \|_p < \epsilon.$$

This shows the sequence in a Cauchy sequence. ■

We will now show the $\mathcal{L}_p(X, \mathcal{S}, \mu)$ is a Banach space.

Theorem 11.1.9. $\mathcal{L}_p$ Is A Banach Space

Let $1 \leq p < \infty$. Then $\mathcal{L}_p(X, \mathcal{S}, \mu)$ is complete with respect to the norm $\| \cdot \|_p$.

Proof. *Let $[\boldsymbol{f_n}]$ be a Cauchy sequence. These are the steps of the proof.*

*(***Step 1***): we find a subsequence $([\boldsymbol{g_k}])$ so that for all k,*

$$\int |g_{k+1} - g_k|^p \, d\mu \quad < \quad (1/2^k)^p \tag{$\boldsymbol{\alpha}$}$$

*(***Step 2***): Define the function g by*

$$g(x) \quad = \quad g_1(x) + \sum_{k=1}^{\infty} |g_{k+1}(x) - g_k(x)|. \tag{$\boldsymbol{\beta}$}$$

We show that g satisfies

$$\| g \|_p \quad \leq \quad \| g_1 \|_p \; + 1. \tag{$\boldsymbol{\gamma}$}$$

This implies that g, defined by Equation $\boldsymbol{\alpha}$, converges and is finite a.e.

*(***Step 3***): Then, we show*

$$f(x) \quad = \quad g_1(x) + \sum_{k=1}^{\infty} \Big(g_{k+1}(x) - g_k(x) \Big)$$

is defined a.e. and is in $L_p(X, \mathcal{S}, \mu)$. This is our candidate for the convergence of the Cauchy sequence.

*(***Step 4***): We show g_k converge to f in $\| \cdot \|_p$.*

*(***Step 5***): We show $[\boldsymbol{f_n}]$ converges to $[\boldsymbol{f}]$ in $\| \cdot \|_p$. This last step will complete the proof of completeness. Now to the proof of these steps.*

*(***Proof Step 1***): For $\epsilon = (1/2)$, since $[\boldsymbol{f_n}]$ is a Cauchy sequence, there is a positive integer N_1 so that $n, m > N_1$ implies*

$$\int |f_n - f_m|^p \, d\mu \; < \; (1/2).$$

Note we use representative $f_n \in [\boldsymbol{f_n}]$ for simplicity of exposition since the norms are independent of choice of representatives. Define $g_1 = f_{N_1+1}$.

Next, for $\epsilon = (1/2)^2$, there is a positive integer N_2, which we can always choose so that $N_2 > N_1$, so that $n, m > N_2$ implies

$$\int |f_n - f_m|^p \, d\mu \; < \; \Big(1/(2^2)\Big)^p.$$

Let $g_2 = f_{N_2+1}$. Then, again by our choice of indices, we have

$$\int |g_2 - g_1|^p \, d\mu \; < \; (1/2)^p.$$

The next step is similar. For $\epsilon = (1/2)^3$, there is a positive integer N_3, which we can always choose so that $N_3 > N_2$, so that $n, m > N_3$ implies

$$\int |f_n - f_m|^p \, d\mu \; < \; \Big(1/(2^3)\Big)^p.$$

Let $g_3 = f_{N_3+1}$. Then, we have

$$\int |g_3 - g_2|^p \, d\mu \; < \; \Big(1/(2^2)\Big)^p.$$

An induction argument thus shows that there is a subsequence $[\boldsymbol{g_k}]$ that satisfies

$$\int |g_{k+1} - g_k|^p \, d\mu \; < \; \Big(1/(2^k)\Big)^p.$$

for all $k \geq 1$. This establishes Equation $\boldsymbol{\alpha}$.

(**Proof Step 2**): *Define the non negative sequence (h_n) by*

$$h_n(x) \; = \; |g_1(x)| \; + \; \sum_{k=1}^{n} |g_{k+1}(x) - g_k(x)|.$$

In this definition, there is the usual messiness of where all the differences are defined. Let's clear that up. Each pair $(g_k, g_{k+1}$ has a potential set E_k of measure zero where the subtraction is not defined. Thus, we need to throw away the set $E = \cup_k E_k$ which also has measure 0. Thus, it is clear that all of the h_n are defined on E^C. Now they may take on the value ∞, but that is acceptable. We see $h_n^p \uparrow g^p$ on E^C. Apply Fatou's Lemma to (h_n). We find

$$\int \Big(\liminf \, h_n^p \, I_{E^C}\Big) \, d\mu \; \leq \; \liminf \int \, h_n^p \, I_{E^C} \, d\mu.$$

But, $\liminf \, h_n^p = g^p$ and so

$$\int \, g^p \, I_{E^C} \, d\mu \; \leq \; \liminf \int \, h_n^p \, I_{E^C} \, d\mu.$$

The p^{th} root function is continuous and so

$$\lim_n \Big(\int \, h_n^p \, I_{E^C} \, d\mu\Big)^{1/p} \; = \; \Big(\lim_n \int \, h_n^p \, I_{E^C} \, d\mu\Big)^{1/p}.$$

Then, since the p^{th} root function is increasing, we have

$$\left(\int g^p \, I_{E^C} \, d\mu\right)^{1/p} \leq \lim_n \left(\int h_n^p \, I_{E^C} \, d\mu\right)^{1/p}.$$

Next, applying Minkowski's Inequality to a finite sum, we obtain

$$\begin{aligned}\left(\int h_n^p \, I_{E^C}\right)^{1/p} &= \left(\int \left(|g_1| + \sum_{k=1}^{n} |g_{k+1} - g_k|\right) I_{E^C}\right)^{1/p} \\ &\leq \| g_1 I_{E^C} \|_p + \sum_{k=1}^{n} \| (g_{k+1} - g_k) I_{E^C} \|_p .\end{aligned}$$

Since the finite sum on the left is monotonic increasing, we have immediately that the series

$$\sum_{k=1}^{\infty} \| (g_{k+1} - g_k) I_{E^C} \|_p$$

is a well defined extended real - valued number. Thus, we have

$$\left(\int h_n^p \, I_{E^C}\right)^{1/p} \leq \| g_1 I_{E^C} \|_p + \sum_{k=1}^{\infty} \| (g_{k+1} - g_k) I_{E^C} \|_p .$$

By Equation $\boldsymbol{\alpha}$*, we also know that*

$$\sum_{k=1}^{\infty} \| (g_{k+1} - g_k) I_{E^C} \|_p \leq \sum_{k=1}^{\infty} 1/(2)^k = 1.$$

Hence, we can actually say

$$\left(\int g \, I_{E^C}\right)^{1/p} \leq \| g_1 I_{E^C} \|_p + 1$$

We conclude $g \, I_{E^C}$ *is in* $L_p(X, \mathcal{S}, \mu)$*. Further, since if* $F = \{x \,|\, g(x) I_{E^C}(x) = \infty\}$*, then we know* F *has measure* 0*. Hence,* $g \, I_{E^C \cap F^C}$ *is finite. This completes Step 2.*

(**Proof Step 3**)*: Now define the function* f *by*

$$f(x) = \begin{cases} g_1(x) + \sum_{k=1}^{\infty} \Big(g_{k+1}(x) - g_k(x)\Big), & x \in E^C \cap F^C \\ 0 & x \in E \cup F. \end{cases}$$

Note, for $x \in E^C \cap F^C$*,*

$$\begin{aligned}|g_k| &= |g_1 + (g_2 - g_1) + (g_3 - g_2) + \ldots + (g_k - g_{k-1}| \\ &\leq |g_1| + \sum_{i=1}^{k} |g_{k+1} - g_k| = h_k.\end{aligned}$$

However, we already seen that on this set $h_k \uparrow g$. Hence, we can say

$$|g_k| \quad \leq \quad g.$$

This tells us that the partial sum expansion of g_k converges absolutely on $E^C \cap F^C$ and thus, g_k converges to g. But $g = f$ on this set, so we have shown that g_k converges to f a.e. We can now apply the Lebesgue Dominated Convergence Theorem to say

$$\lim_n \int g_n \, d\mu \quad = \quad \int f \, d\mu.$$

Since $|g_k| \leq g$ for all k, it follows $|f|^p \leq |g|^p$. Since g is p summable, we have established that f is in $L_p(X, \mathcal{S}, \mu)$.

(**Proof Step 4**): *Now we show g_k converges to f in $\| \cdot \|_p$. To see this, let $z_k = f - g_k$ on $E^C \cap F^C$. From the definition of f, we can write this as $\sum_{j=k}^{\infty} (g_{j+1} - g_j)$. The rest of the argument is very similar to the one used in Step 2. Consider the partial sums of this convergent series*

$$z_k^n \;=\; \sum_{j=k}^{n} |g_{j+1} - gj|.$$

Minkowski's Inequality then gives for all n,

$$\| \, z_k^n \, \|_p \leq \sum_{j=k}^{n} \| \, g_{j+1} - gj \, \|_p \; .$$

Using Equation $\boldsymbol{\alpha}$, it follows that the right hand side is bounded above by $\sum_{j=k}^{n} 1/2^j$ which sums to $1/2^{n-1}$. Now apply Fatou's Lemma to find

$$\int \liminf \, |z_k^n|^p \;\leq\; \liminf \int \, |z_k^n|^p$$

or

$$\int \, |z_k|^p \;\leq\; \liminf \int \, |z_k^n|^p.$$

The continuity and increasing nature of the p^{th} root then give us

$$\left(\int \, |z_k|^p \right)^{1/p} \;\leq\; \liminf \left(\int \, |z_k^n|^p \right)^{1/p} \;\leq\; \liminf \, (1/2^{n-1}) \;=\; 0.$$

Thus, $\| \, f - g_k \, \| \to 0$.

(**Proof Step 5**): *Finally, given $\epsilon > 0$, since $[\boldsymbol{f_n}]$ is a Cauchy sequence, there is an N so that*

$$\| \, f_n - f_m \, \|_p < \epsilon/2, \; \forall \, n, m \; > \; N.$$

Since $[\boldsymbol{g_k}]$ is a subsequence of $[\boldsymbol{f_n}]$, there is a K_1 so that if $k > K_1$, we have

$$\| f_m - g_k \|_p < \epsilon/2, \; \forall\, m \;>\; N, \; k \;>\; K_1.$$

Also, since $g_k \rightarrow f$ in p - norm, there is a K_2 so that

$$\| g_k - f \|_p < \epsilon/2, \; \forall\, k \;>\; K_2.$$

We conclude for any given $k > \max(K_1, K_2)$, we have

$$\| f_m - f \|_p \leq \| f_m - g_k \|_p \; + \; \| g_k - f \|_p \; < \; \epsilon$$

if $m > N4$. Thus, $[\boldsymbol{f_n}] \rightarrow [\boldsymbol{f}]$ in p - norm. ∎

The proof of the theorem above has buried in it a powerful result. We state this below.

Theorem 11.1.10. Sequences That Converge In p - Norm Possess Subsequences Converging Pointwise a.e.
Let $1 \leq p < \infty$. Let $([\boldsymbol{f_n}])$ be a sequence in $\mathcal{L}_p(X, \mathcal{S}, \mu)$ which converges in norm to $[\boldsymbol{f}]$ in $\mathcal{L}_p(X, \mathcal{S}, \mu)$. Then, there is a subsequence $([\boldsymbol{f_n^1}]$ of $([\boldsymbol{f_n}])$ which converges pointwise a.e. to f.

Proof. *The sequence we seek is the sequence (g_n) as defined in the proof of Theorem 11.1.9; see the discussion for the proof of Step (3).* ∎

11.2 The World Of Counting Measure

Let's see what the previous material means when we use counting measure, μ_C, on the set of natural numbers $\mathbb{N}$. In this case, the sigma - algebra is the power set of $\mathbb{N}$. Note if $f : \mathbb{N} \rightarrow \overline{\Re}$, then f is identified with a sequence of extended real - valued numbers, (a_n) so that $f(n) = a_n$. It is therefore possible for $f(n) \; \infty$ or $f(n) = -\infty$ for some n. Let

$$\phi_N(n) \;=\; \begin{cases} |f(n)|, & 1 \;\leq\; n \;\leq\; N \\ 0, & n \;>\; N \end{cases}$$

Then, $\phi_N \uparrow f$ and so by the Monotone Convergence Theorem,

$$\int |f| \, d\mu_C \quad = \quad \lim_N \int \phi_N(n) \, d\mu_C.$$

Now the simple functions ϕ_N are not in their standard representation. Let $\{c_1, \ldots, c_M\}$ be the distinct elements of $\{|a_1|, \ldots, |a_N|\}$. Then we can write

$$\phi_N \;=\; \sum_{i=1}^{M} c_i \, I_{E_i},$$

where E_I is the pre-image of each distinct element c_i. The sets E_i are clearly disjoint by construction. It is a straightforward matter to see that

$$\int \phi_N \, d\mu_C \; = \; \sum_{i=1}^{M} c_i \, \mu_C E_i \; = \; \sum_{i=1}^{N} |a_i|.$$

Thus, we have

$$\int |f| \, d\mu_C \quad = \quad \lim_{N} \sum_{i=1}^{N} |f(i)|.$$

Since all the terms $|f(i)|$ are non negative, we see the sequence of partial sums converges to some extended real - valued number (possibly ∞). For counting measure, the only set of measure 0 is $\emptyset$, so measurable functions can not differ on a set of measure 0 in this case. We see for $1 \leq p < infty$,

$$L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_C) \; = \; \mathcal{L}_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_C).$$

Further,

$$L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_C) \; = \; \{ \text{ sequences } \; (a_n) \, | \, \sum_{i=1}^{\infty} |a_i|^p \; \text{ converges } \}.$$

We typically use the notation

$$\ell_p \; = \; L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_C) \; = \; \{ \text{ sequences } \; (a_n) \, | \, \sum_{i=1}^{\infty} |a_i|^p \; \text{ converges } \}.$$

and we call this a sequence space. Note in all cases, summability implies the sequence involved must be finite everywhere.

In this context, Hölder's Inequality becomes:

Theorem 11.2.1. Hölder's Inequality: Sequence Spaces

Let $1 < p < \infty$ and q be the index conjugate to p. Let $(a_n) \in \ell_p$ and $(b_n) \in ell_q$. Then $(a_n \, b_n) \in \ell_1$ and

$$\sum_n |a_n \, b_n| \; \leq \; \left(\sum_n |a_n|^p \right)^{1/p} \left(\sum_n |b_n|^q \right)^{1/q}.$$

and Minkowski's Inequality becomes

Theorem 11.2.2. Minkowski's Inequality: Sequence Spaces

Let $1 \leq p < \infty$ and let $(a_n), (b_n) \in \ell_p$. Then $(a_n + b_n)$ is in ℓ_p and

$$\left(\sum_n |a_n \, + \, b_n|^p \right)^{1/p} \; \leq \; \left(\sum_n |a_n|^p \right)^{1/p} \; + \; \left(\sum_n |b_n|^p \right)^{1/p}.$$

Finally, the special case of $p = q = 2$ should be mentioned. The sequence space version of the resulting Hölder's Inequality Cauchy - Schwartz Inequality has this form:

Theorem 11.2.3. Cauchy - Bunyakovskiĭ - Schwartz Inequality: Sequence Spaces
Let $(a_n), (b_n) \in \ell_2$. Then $(a_n\, b_n) \in \ell_1$ and

$$\sum_n |a_n\, b_n| \;\leq\; \left(\sum_n |a_n|^2\right)^{1/2} \left(\sum_n |b_n|^2\right)^{1/2}.$$

11.3 Equivalence Classes of Essentially Bounded Functions

There is one more space to define. This will be the analogue of the space of bounded functions we use in the definition of the Riemann Integral.

Definition 11.3.1. Essentially Bounded Functions
Let $(X, \mathcal{S}, \mu)$ be a measure space and let f be measurable. If E is a set of measure 0, let

$$\xi(E) \;=\; \sup_{x \in E^C} |f(x))$$

and

$$\rho_\infty(f) \;=\; \inf\{\, \xi(E) \mid E \in \mathcal{S},\ \mu(E) = 0\}.$$

If $\rho_\infty(f)$ is finite, we say f is an essentially bounded function.

There are then two more spaces to consider:

Definition 11.3.2. The Spaces of Essentially Bounded Functions
Let $(X, \mathcal{S}, \mu)$ be a measure space. Then we define

$$L_\infty(X, \mathcal{S}, \mu) \;=\; \{f : X \to \overline{\Re} \mid f \in M(X, \mathcal{S}),\ \rho_\infty(f) < \infty\}.$$

and defining equivalence classes using a.e. equivalence,

$$\mathcal{L}_\infty(X, \mathcal{S}, \mu) \;=\; \{[\boldsymbol{f}] \mid \rho_\infty(f) < \infty\}.$$

There is an equivalent way of characterizing an essentially bounded function. This requires another definition.

Theorem 11.3.1. Alternate Characterization Of Essentially Bounded Functions
Let $(X, \mathcal{S}, \mu)$ be a measure space and f be a measurable function. Define $q_\infty(f)$ by

$$q_\infty(f) \;=\; \inf\{\, a \mid \mu(\{x \mid |f(x)| > a\}) \;=\; 0\}.$$

Then, $\rho_\infty(f) = q_\infty(f)$.

Proof. *Let $E_a = \{x \mid |f(x)| > a\}$. If a is a number so that $\mu(E_a) = 0$, then for any other measurable set A with measure 0, we have*

$$A^C \;=\; A^C \cap E_a \,\cup\, A^C \cap E_a^C.$$

Thus,

$$\sup_{A^C} |f| \;\geq\; \sup_{A^C \cap E_a} |f| \;\geq\; a.$$

because if $x \in A^C \cap E_a$, then $|f(x) > a$. Since we can do this for such a, it follows that

$$\sup_{A^C} |f| \;\geq\; q_\infty(f).$$

Further, since the measurable set A with measure zero is arbitrary, we must have

$$\rho_\infty(f) \;\geq\; q_\infty(f).$$

Next, we prove the reverse inequality. If $\mu(E_a) = 0$, then by the definition of $\rho_\infty(f)$, we have

$$\rho_\infty(f) \;\leq\; \sup_{E_a^C} |f| \;=\; \sup_{|f(x)| \leq a} |f(x)| \;\leq\; a.$$

But this is true for all such a. Thus, $\rho_\infty(f)$ is a lower bound for the set $\{a | \mu(E_a) = 0\}$ and we can say

$$\rho_\infty(f) \;\leq\; q_\infty(f).$$

■

We need to know that if two functions are equivalent with respect to the measure μ, then their ρ_∞ values agree.

Lemma 11.3.2. Essentially Bounded Functions That Are Equivalent Have The Same Essential Bound

Let $(X, \mathcal{S}, \mu)$ be a measure space and f and g be a equivalent measurable functions such that $\rho(f)$ is finite. Then $\rho(g) = \rho(f)$.

Proof. *Let E be the set of points where f and g are not equal. Then $\mu(E) = 0$. Now,*

$$0 \;\leq\; \mu\Big((|g(x)| > a) \;\cap\; E \Big) \;\leq\; \mu(E) \;=\; 0.$$

Thus,

$$\begin{aligned} \mu\Big((|g(x)| > a) \Big) \;&=\; \mu\Big((|g(x)| > a) \;\cap\; E \Big) \;+\; \mu\Big((|g(x)| > a) \;\cap\; E^C \Big) \\ &=\; \mu\Big((|g(x)| > a) \;\cap\; E^C \Big). \end{aligned}$$

But on E^C, f and g match, so we have

$$\mu\Big((|g(x)| > a) \Big) \;=\; \mu\Big((|f(x)| > a) \;\cap\; E^C \Big) \;=\; \mu\Big((|f(x)| > a) \Big),$$

by the same sort of argument we used on $\mu\Big((|g(x)| > a)\Big)$. *Hence, if* $\mu\Big((|f(x)| > a)\Big) = 0$, *then* $\mu\Big((|g(x)| > a)\Big) = 0$ *as well. This immediately implies* $q_\infty(g) = q_\infty(f)$. *The result then follows because* $q_\infty = \rho_i nfty$. ■

Finally, we can show that essentially bounded functions are bounded above by their essential bound a.e.

Lemma 11.3.3. Essentially Bounded Functions Bounded Above By Their Essential Bound a.e

Let $(X, \mathcal{S}, \mu)$ *be a measure space and* f *be a measurable functions such that* $\rho(f)$ *is finite. Then* $|f(x)| \leq \rho(f)$ *a.e.*

Proof. *Let* $E = (|f(x) > \rho_\infty(f))$. *It is easy to see that*

$$E \;=\; \bigcup_{k=1}^{\infty} \Bigg(|f(x)| > \rho_\infty(f) + 1/k \Bigg).$$

If you look at how q_∞ *is defined, if* $\mu(|f(x)| > \rho_\infty(f) + 1/k) > 0$, *that would force* $q_\infty(f) = \rho_\infty(f) \geq \rho_\infty(f) + 1/k$ *which is not possible. Hence,* $\mu(|f(x)| > \rho_\infty(f) + 1/k) = 0$ *for all* k. *This means* E *has measure* 0 *also. It is then clear from the definition of the set* E *that* $|f(x)| \leq \rho_\infty(f)$ *on* E^C. ■

We can now prove that $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$ is a vector space with norm $\| \, [\boldsymbol{f}] \, \|_\infty = \rho_\infty(f)$.

Theorem 11.3.4. The L_∞ Semi-Norm

Let $(X, \mathcal{S}, \mu)$ *be a measure space Define* $\| \cdot \|_\infty$ *on* $L_\infty(X, \mathcal{S}, \mu)$ *by*

$$\| \; f \; \|_\infty \quad = \quad \rho_\infty(g),$$

where g *is any representative of* $[\boldsymbol{f}]$. *Then,* $\| \cdot \|_\infty$ *is a semi-norm.*

Proof. *We show* $\rho_\infty(\cdot)$ *satisfies all the properties of a norm except N2 and hence it is a semi - norm.*
(N1): It is clear the N1 is satisfied because $\rho_\infty(\cdot)$ *is always non negative.*
(N2): Let 0_X *is the function defined to be* 0 *for all* x *and let* $E_a = \{x \mid |0_X(x)| > a\}$. *It is clear* $E_a = \emptyset$ *for all* $a > 0$. *Thus, since* $\rho_\infty = q_\infty$,

$$q_\infty(0_X) \quad = \quad \inf\{\, a \mid \mu(E_a) \; = \; 0\} \; = \; 0.$$

However, if $q_\infty(f) = 0$, *let* $F_n = (|f_n(x)| > 1/n)$. *Then, by definition of* $q_\infty(f)$, *it follows that* $\mu(F_n) = 0$ *and* $|f(x)| \leq 1/n$ *on the complement* F_n^C. *Let* $F = \cup F_n$. *Then,* $\mu(F) = 0$ *and*

$$F^C \;=\; \bigcap_n F_n^C \;=\; \bigcap_n \Bigg(|f(x)| \leq 1/n \Bigg) \;=\; \Bigg(f(x) = 0).$$

Thus, f *is* 0 *on* F^C *and non zero on* F *which has measure* 0. *All that we can say then is that* $f = 0$ *a.e. and hence,* $\| \cdot \|_\infty$ *does not satisfy N2.*

(N3): If α is 0, the result is clear. If α is a non zero number, then

$$\begin{aligned} q_\infty(\alpha f) &= \inf \{ \, a \mid \mu(\{x \mid |\alpha\, f(x)| > a\}) \;=\; 0\} \\ &= \inf \{ \, a \mid \mu(\{x \mid |f(x)| > a/\alpha\}) \;=\; 0\}. \end{aligned}$$

Let $\beta = a/\alpha$ and we have

$$\begin{aligned} q_\infty(\alpha f) &= \inf \{ \, \alpha\, \beta \mid \mu(\{x \mid |f(x)| > \beta\}) \;=\; 0\} \\ &= \alpha\, \inf \{ \, \beta \mid \mu(\{x \mid |f(x)| > \beta\}) \;=\; 0\} \\ &= \alpha\, q_\infty(\alpha f). \end{aligned}$$

(N4): Now let f and g be in $L_\infty(X, \mathcal{S}, \mu)$ with the sum $f+g$ defined in the usual way on E_{fg}^C with $\mu(E_{fg}) = 0$. Note on E_{fg} itself, $f(x) + g(x) = 0$, so the sum is bounded above by $\rho_\infty(f) + \rho_\infty(g)$ there. Now by Lemma 11.3.3, there are sets F and G of measure 0 so that

$$\begin{aligned} |f(x)| &\leq \rho_\infty(f), \, \forall x \in F^C, \\ |g(x)| &\leq \rho_\infty(g), \, \forall x \in G^C. \end{aligned}$$

Thus,

$$|f(x) + g(x)| \;\leq\; \rho_\infty(f) \;+\; \rho_\infty(g), \, \forall x \in F^C \;\cap\; G^C,$$

Thus, the measure of the set of points where $|f(x) + g(x)| > \rho_\infty(f) \;+\; \rho_\infty(g)$ is zero as $\mu(F \cup G) = 0$. By definition of q_∞, it then follows that

$$q_\infty(f+g) \;\leq\; \rho_\infty(f) \;+\; \rho_\infty(g).$$

which implies the result. ∎

Theorem 11.3.5. $\mathcal{L}_\infty$ Is A Normed Linear Space

Then $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$ is a vector space over $\Re$ with scalar multiplication and object addition defined as

$$\begin{aligned} \alpha\,[\boldsymbol{f}] &= [\boldsymbol{\alpha}\; \boldsymbol{f}]\; \forall\, [\boldsymbol{f}] \\ [\boldsymbol{f}] \,+\, [\boldsymbol{g}] &= [\boldsymbol{f} \;+\; \boldsymbol{g}], \forall [\boldsymbol{f}] \;\; \textit{and} \;\; [\boldsymbol{g}]. \end{aligned}$$

Further, $\|\, [\boldsymbol{f}] \,\|_\infty$ defined by

$$\|\, [\boldsymbol{f}] \,\|_\infty \;=\; \rho_\infty(g),$$

for any representative g of $[\boldsymbol{f}]$ is a norm on $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$.

Proof. *The argument that the scalar multiplication and addition of equivalence classes is the same as the one we used in the proof of Theorem 11.1.5 and so we will not repeat it here. From Lemma 11.3.2 we know that any two functions which are equivalent a.e. will have the same value for ρ_∞ and so $\|\, [\boldsymbol{f}] \,\|_\infty$*

is independent of the choice of representative from $[\boldsymbol{f}]$. *The proofs that properties N1, N3 and N4 hold follow immediately from the fact that they hold for representatives of equivalence classes. It remains to show that if* $\| \, [\boldsymbol{f}] \, \|_\infty = 0$, *then* $[\boldsymbol{f}] = [\boldsymbol{0_X}]$ *where* 0_X *is the zero function on* X. *However, we have already established in the proof of Theorem 11.3.4 that such an* f *is* 0 *a.e. This tells us* $f \in [\boldsymbol{0_X}]$; *thus,* $[\boldsymbol{f}] = [\boldsymbol{0_X}]$. ■

Theorem 11.3.6. $\mathcal{L}_\infty$ **Is A Banach Space**

Then $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$ *is complete with respect to the norm* $\| \cdot \|_\infty$.

Proof. *Let* $([\boldsymbol{f_n}]$ *be a Cauchy sequence of objects in* $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$. *Now everything is independent of the choice of representative of an equivalence class, so for convenience, we will use as our representatives the functions* f_n *themselves. Then, by Lemma 11.3.3, there are sets* E_n *of measure* 0 *so that*

$$|f_n(x)| \;\le\; \rho_\infty(f_n), \; \forall \, x \in \; E_n^C.$$

Also, there are sets F_{nm} *of measure* 0 *so that*

$$|f_n(x) \;-\; f_m(x)| \;\le\; \rho_\infty(f_n \;-\; f_m), \; \forall \, x \in \; F_{nm}^C.$$

Hence, both of the equations above hold on

$$U \;=\; \bigcap_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \Big(E_n^C \cap F_{nm}^C \Big).$$

We then use De Morgan's Laws to rewrite U *as follows:*

$$\begin{aligned} U \;&=\; \bigcap_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \Big(E_n \cup F_{nm} \Big)^C \\ &=\; \bigcap_{m=1}^{\infty} \bigg(\bigcup_{n=1}^{\infty} \Big(E_n \cup F_{nm} \Big) \bigg)^C \\ &=\; \bigg(\bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \Big(E_n \cup F_{nm} \Big) \bigg)^C. \end{aligned}$$

Clearly, the measure of U *is* 0 *and*

$$|f_n(x) \;-\; f_m(x)| \;\;\le\;\; \rho_\infty(f_n - f_m), \; \forall \, x \; \in \; U^C. \tag{$*$}$$

Now since $([\boldsymbol{f_n}]$ *is a Cauchy sequence with respect to* $\| \cdot \|_\infty$, *given* $\epsilon > 0$, *there is a positive integer* N *so that*

$$|f_n(x) \;-\; f_m(x)| \;\;\le\;\; \rho_\infty(f_n - f_m) \;<\; \epsilon/4, \; \forall \, x \; \in \; U^C, \; \forall \, n, \; m \; > \; N. \tag{$**$}$$

Equation $**$ *implies that at each* x *in* U^C, *the sequence* $(f_n(x))$ *is a Cauchy sequence of real numbers. By the completeness of* $\Re$, *it then follows that* $\lim_n \; f_n(x)$ *exists on* U^C. *Define the function* $f : X \to \Re$

by

$$f(x) \quad = \quad \begin{cases} \lim_n f_n(x), & x \in U^C, \\ 0 & x \in U. \end{cases}$$

From Equation $**$, *we have that*

$$\lim_n |f_n(x) - f_m(x)| \leq \epsilon/4, \forall x \in U^C, \forall m > N.$$

As usual, since the absolute value function is continuous, we can let the limit operation pass into the absolute value function to obtain

$$|f(x) - f_m(x)| \leq \epsilon/4, \forall x \in U^C, \forall m > N. \qquad (***)$$

From the backwards triangle inequality, we then find

$$|f(x)| \quad \leq \quad \epsilon/4 + |f_m(x)| < \epsilon/4 + \rho_\infty(f_m), \forall x \in U^C, \forall m > N.$$

Now fix $M > N + 1$. *Then*

$$|f(x)| \quad < \quad \epsilon/2 + \rho_\infty(f_M), \forall x \in U^C.$$

Since the measure of the set $(|f(x)| > \epsilon/4 + \rho_\infty(f_M)$ *is* 0, *from the definition of* $q_\infty(f)$, *it then follows that*

$$q_\infty(f) \leq \epsilon/4 + \rho_\infty(f_M)$$

which tells us that f *is essentially bounded.*

It remains to show that $[\boldsymbol{f_n}]$ *converges to* $[\boldsymbol{f}]$ *in norm. Note that Equation* $***$ *implies that* (f_n) *converges uniformly on* U^C. *Further,the measure of the set* $(|f_n(x) - f(x)| > \epsilon/4)$ *is* 0. *Thus, we can conclude*

$$q_\infty(f - f_n) \leq \epsilon/4 < \epsilon, \forall n > N.$$

This shows the desired convergence in norm.

Thus, we have shown $\mathcal{L}_\infty(X, \mathcal{S}, \mu)$ *is complete.* ■

From the proofs above, we see Minkowski's Inequality holds for the case $p = \infty$ because $\| \cdot \|_\infty$ is a norm. Finally, we can complete the last case of Hölder's Inequality: the case of the conjugate indices $p = 1$ and $q = \infty$. We obtain

Theorem 11.3.7. Hölder's Inequality: $p = 1$

Let $p = 1$ *and* $q = \infty$ *be the index conjugate to* 1. *Let* $[\boldsymbol{f}] \in \mathcal{L}_1(X, \mathcal{S}, \mu)$ *and* $[\boldsymbol{g}] \in \mathcal{L}_\infty(X, \mathcal{S}, \mu)$. *Then* $[\boldsymbol{f}\, \boldsymbol{g}] \in \mathcal{L}_1(X, \mathcal{S}, \mu)$ *and*

$$\int |fg| \, d\mu \leq \| [\boldsymbol{f}] \|_1 \, \| [\boldsymbol{g}] \|_\infty .$$

Proof. *it is enough to prove this result for the representatives of the equivalence classes $f \in [\boldsymbol{f}]$ and $g \in [\boldsymbol{g}]$. We know the product fg is measurable. It remains to show that fg is summable. Since g is essentially bounded, by Lemma 11.3.3, there is a sets E of measure 0 so that*

$$|g(x)| \leq \rho_\infty(g), \forall x \in E^C.$$

Thus, $|f(x)\, g(x)| \leq |f(x)|\, \rho_\infty(g)$ a.e. and since the right hand side is summable, by Theorem 10.6.1, we see fg is also summable and

$$\int |f\, g|\, d\mu \leq \int |f|\, \rho_\infty(g)\, d\mu = \rho_\infty(g) \int |f|\, d\mu$$

■

11.4 The Hilbert Space L_2

The space $\mathcal{L}_2(X, \mathcal{S}, \mu)$ is a Normed linear space with norm $\| [\cdot] \|_2$. This space is also an inner product space which is complete. Such a space is called a Hilbert space.

Definition 11.4.1. Inner Product Space

Let X be a non empty vector space over $\Re$. Let $\omega\, X \times X \rightarrow \Re$ be a mapping which satisfies

IP1:

$$\omega(x + y, z) = \omega(x, z) + \omega(y, x), \forall x, y, z \in X,$$

IP2:

$$\omega(\alpha\, x, y) = \alpha\, \omega(x, y), \forall \alpha \in \Re, \forall x, y \in X,$$

IP3:

$$\omega(x, y) = \omega(y, x), \forall x, y \in X,$$

IP4:

$$\omega(x, x) \geq 0, \forall x, \in X, \text{ and } \omega(x, x) = 0 \Leftrightarrow x = 0.$$

Such a mapping is called an real inner product on the real vector space X. It is easy to define a similar mapping on complex vector spaces, but we will not do that here. We typically use the symbol $< \cdot, \cdot >$ to denote the value $\omega(\cdot, \cdot)$.

There is much more we could say on this subject, but instead we will focus on how we can define an inner product on $\mathcal{L}_2(X, \mathcal{S}, \mu)$.

Theorem 11.4.1. The Inner Product on The Space of Square Summable Equivalence Classes
For brevity, let $\mathcal{L}_2$ denote $\mathcal{L}_2(X, \mathcal{S}, \mu)$. The mapping $< \cdot, \cdot >$ on $\mathcal{L}_2 \times \mathcal{L}_2$ defined by

$$< [\boldsymbol{f}], [\boldsymbol{g}] > \; = \; \int u\, v\, d\mu, \; \forall\, u \in [\boldsymbol{f}], \; v \in [\boldsymbol{g}]$$

is an inner product on $\mathcal{L}_2$. Moreover, it induces the norm $\| [\cdot] \|_2$ by

$$\begin{aligned} \| [\boldsymbol{f}] \|_2 &= \sqrt{\int |f|^2 \, d\mu} \\ &= \sqrt{< [\boldsymbol{f}], [\boldsymbol{f}] >}. \end{aligned}$$

Proof. *The proof of these assertions is immediate as we have already shown $\| \cdot \|_2$ is a norm and the verification of properties IP1 to IP4 is straightforward.* ∎

Finally, from our general $\mathcal{L}_p$ results, we know $\mathcal{L}_2$ is complete. However, for the record, we state this as a theorem.

Theorem 11.4.2. The Space of Square Summable Equivalence Classes Is A Hilbert Space
For brevity, let $\mathcal{L}_2$ denote $\mathcal{L}_2(X, \mathcal{S}, \mu)$. Then $\mathcal{L}_2$ is complete with respect to the norm induced by the inner product $< [\cdot], [\cdot] >$. The inner product space $(\mathcal{L}_2, < \cdot, \cdot >)$ is often denoted by the symbol $\mathcal{H}$.

11.5 Homework

Exercise 11.5.1. *Let $(X, \mathcal{S}, \mu)$ be a measure space. Let f be in $\mathcal{L}_p(X, \mathcal{S}\mu)$ for $1 \leq p < \infty$. Let $E = \{x \mid |f(x)| \neq 0\}$. Prove E is σ - finite.*

Exercise 11.5.2. *Let $(X, \mathcal{S}, \mu)$ be a finite measure space. If f is measurable, let $E_n = \{x \mid n - 1 \leq |f(x)| < n\}$. Prove f is in $\mathcal{L}_1(X, \mathcal{S}\mu)$ if and only if $\sum_{n=1}^{\infty} n\mu(E_n) < \infty$.*
More generally, prove f is in $\mathcal{L}_p(X, \mathcal{S}\mu)$, $1 \leq p < \infty$, if and only if $\sum_{n=1}^{\infty} n^p \mu(E_n) < \infty$.

Chapter 12

Constructing Measures

Although you now know quite a bit about measures, measurable functions, associated integration and the like, you still do not have many concrete and truly interesting measures to work with. In this chapter, you will learn how to construct interesting measures using some simple procedures. A very good reference for this material is (Bruckner et al. (1) 1997) . Another good source is (Taylor (7) 1985) . We begin with a definition.

12.1 Measures From Outer Measures

Definition 12.1.1. Outer Measure
Let X be a non empty set and let μ^* be an extended real valued mapping defined on all subsets of X that satisfies

(i): $\mu^*(\emptyset) = 0$.

(ii): If A and B are subsets of X with $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$.

(iii): If (A_n) is a sequence of subsets of X, then $\mu^*(\cup_n A_n) \leq \sum_n \mu^*(A_n)$.

Such a mapping is an outer measure on X and condition (iii) is called the countable subadditivity (CSA) condition if the sets are disjoint.

Comment 12.1.1. *Since $\emptyset \subseteq A$ for all A in X, condition (ii) tells us $\mu^*(\emptyset) \leq \mu^*(A)$. Hence, by condition (i), we have $\mu^*(A) \geq 0$ always. Thus, the outer measure is non negative.*

The outer measure is defined on all the subsets of X. In Chapter 10, we defined the notion of a measure on a σ - algebra of subsets of X. Look back at Definition 10.0.1 again. Recall, the mapping $\mu : \mathcal{S} \rightarrow \overline{\Re}$ is a measure on $\mathcal{S}$ if

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

(ii): $\mu(E) \geq 0$, for all $E \in \mathcal{S}$, (iii): μ is countably additive on $\mathcal{S}$; i.e. if $(E_n) \subseteq \mathcal{S}$ is a countable collection of disjoint sets, then $\mu(\cup_n E_n) = \sum_n \mu(E_n)$.

The third condition says the mapping μ is countably additive and hence, we label this condition as condition (CA). The collection of all subsets of X is the largest σ - algebra of subsets of X, so to convert the outer measure μ^* into a measure, we have to convert the countable subadditivity condition to countable additivity. This is not that easy to do! Now if T and E are any subsets of X, then we know

$$T \;=\; \Big(T \cap E\Big) \bigcup \Big(T \cap E^C\Big).$$

The outer measure μ^* is subadditive on finite disjoint unions and so we always have

$$\mu^*(T) \;\leq\; \mu^*\Big(T \cap E\Big) + \mu^*\Big(T \cap E^C\Big).$$

To have equality, we need to have

$$\mu^*(T) \;\geq\; \mu^*\Big(T \cap E\Big) + \mu^*\Big(T \cap E^C\Big),$$

also. So, as a first set towards the countable additivity condition we need, why don't we look at all subsets E of X that satisfy the condition

$$\mu^*(T) \;\geq\; \mu^*\Big(T \cap E\Big) + \mu^*\Big(T \cap E^C\Big), \; \forall\, T \subseteq X.$$

We don't know how many such sets E there are at this point. But we certainly want finite additivity to hold. Therefore, it seems like a good place to start. This condition is called the Caratheodory Condition.

Definition 12.1.2. Caratheodory Condition

Let μ^* be an outer measure on the non empty set X. A subset E of X satisfies the Caratheodory Condition if for all subsets T,

$$\mu^*(T) \;\geq\; \mu^*\Big(T \cap E\Big) + \mu^*\Big(T \cap E^C\Big).$$

Such a set E is called μ^* measurable. The collection of all μ^* measurable subsets of X will be denoted by $\mathcal{M}$.

We will first prove that the collection of μ^* measurable sets is an algebra of sets.

Definition 12.1.3. Algebra Of Sets

Let X be a non empty set and let $\mathcal{A}$ be a nonempty family of subsets of X. We say $\mathcal{A}$ is an algebra of sets if

(i): $\emptyset$ is in $\mathcal{A}$.

(ii): If A and B are in $\mathcal{A}$, so is $A \cup B$.

(iii): if A is in $\mathcal{A}$, so is $A^C = X \setminus A$.

Theorem 12.1.1. The μ^* Measurable Sets Form An Algebra

Let X be a non empty set, μ^* an outer measure on X and $\mathcal{M}$ be the collection of μ^* measurable subsets of X. Then $\mathcal{M}$ is a algebra.

Proof. *For the empty set,*

$$\begin{aligned} \mu^*\Big(T \cap \emptyset\Big) + \mu^*\Big(T \cap \emptyset^C\Big) &= \mu^*\Big(\emptyset\Big) + \mu^*\Big(T \cap X\Big) \\ &= 0 + \mu^*\Big(T\Big). \end{aligned}$$

Hence $\emptyset$ satisfies the Caratheodory condition and so $\emptyset \in \mathcal{M}$.

Next, if $A \in \mathcal{M}$, we note the Caratheodory condition is symmetric with respect to complementation and so $A^C \in \mathcal{M}$ also.

To show $\mathcal{M}$ is closed under countable unions, we will start with the union of just two sets and then proceed by induction. Let E_1 and E_2 be in $\mathcal{M}$. Let T be in X. Then, since E_1 and E_2 both satisfy Caratheodory's condition, we know

$$\mu^*(T) = \mu^*(T \cap E_1) + \mu^*(T \cap E_1^C) \qquad (\boldsymbol{a})$$

and

$$\mu^*(T) = \mu^*(T \cap E_2) + \mu^*(T \cap E_2^C). \qquad (\boldsymbol{b})$$

In Equation ***b****, let "T" be "$T \cap E_1^C$". This gives*

$$\mu^*(T \cap E_1^C) = \mu^*(T \cap E_1^C \cap E_2) + \mu^*(T \cap E_1^C \cap E_2^C). \qquad (\boldsymbol{c})$$

We also know that

$$T \cap E_1 = T \cap (E_1 \cup E_2) \cap E_1, \; T \cap E_1^C \cap E_2 = T \cap (E_1 \cup E_2) \cap E_1^C. \qquad (\boldsymbol{d})$$

Now replace the term "$\mu^*(T \cap E_1^C)$" in Equation ***a*** *by the one in Equation* ***c****. This gives*

$$\mu^*(T) = \mu^*(T \cap E_1) + \mu^*(T \cap E_1^C \cap E_2) + \mu^*(T \cap E_1^C \cap E_2^C).$$

Next, replace the sets in the first two terms on the right side in the equation above by what is shown in Equation ***d****. We obtain*

$$\mu^*(T) = \mu^*(T \cap (E_1 \cup E_2) \cap E_1) + \mu^*(T \cap (E_1 \cup E_2) \cap E_1^C) + \mu^*(T \cap E_1^C \cap E_2^C).$$

But E_1 is in $\mathcal{M}$ and so

$$\mu^*(T \cap (E_1 \cup E_2)) = \mu^*(T \cap (E_1 \cup E_2) \cap E_1) + \mu^*(T \cap (E_1 \cup E_2) \cap E_1^C).$$

Using this identity, we then have

$$\begin{aligned}\mu^*(T) &= \mu^*(T \cap (E_1 \cup E_2)) + \mu^*(T \cap E_1^C \cap E_2^C) \\ &= \mu^*(T \cap (E_1 \cup E_2)) + \mu^*(T \cap (E_1 \cup E_2)^C),\end{aligned}$$

using DeMorgan's laws. Since the set T is arbitrary, we have shown $E_1 \cup E_2$ is also in $\mathcal{M}$.

Since, E_1 and E_2 are in $\mathcal{M}$, we now know $E_1^C \cup E_2^C$ is in $\mathcal{M}$ too. But this set is the same as $E_1 \cap E_2$. Thus, $\mathcal{M}$ is closed under intersection.

It then follows that $E_1 \setminus E_2 = E_1 \cap E_2^C$ is in $\mathcal{M}$. So $\mathcal{M}$ is also closed under set differences. Hence, $\mathcal{M}$ is an algebra. ■

Theorem 12.1.2. μ^* **Measurable Sets Properties**
Let X be a non empty set, μ^* an outer measure on X and $\mathcal{M}$ be the collection of μ^* measurable subsets of X. Then, if (E_n) is a countable disjoint sequence from $\mathcal{M}$, $\cup_n E_n$ is in $\mathcal{M}$ and

$$\mu^*(T \cap \cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu^*\Big(T \cap E_i)\Big).$$

for all T in X.

Proof. *Let "T" be "$T \cap (E_1 \cup E_2)$ in the Caratheodory condition of E_2. Then, we have*

$$\mu^*(T \cap (E_1 \cup E_2)) = \mu^*(T \cap (E_1 \cup E_2) \cap E_2) + \mu^*(T \cap (E_1 \cup E_2) \cap E_2^C).$$

This simplifies to

$$\mu^*(T \cap (E_1 \cup E_2)) = \mu^*(T \cap E_2) + \mu^*(T \cap E_1 \cap E_2^C).$$

But E_1 and E_2 are disjoint. Hence, E_1 is contained in E_2^C. Hence, we can further simplify to

$$\mu^*(T \cap (E_1 \cup E_2)) = \mu^*(T \cap E_2) + \mu^*(T \cap E_1).$$

Let's do another step. Since E_3 is in $\mathcal{M}$, we have

$$\begin{aligned}\mu^*(T \cap (E_1 \cup E_2 \cup E_3)) &= \mu^*(T \cap (E_1 \cup E_2 \cup E_3) \cap E_3) \\ &\quad + \mu^*(T \cap (E_1 \cup E_2 \cup E_3) \cap E_3^C).\end{aligned}$$

This can be rewritten as

$$\begin{aligned}\mu^*(T \cap (E_1 \cup E_2 \cup E_3)) &= \mu^*(T \cap E_3) + \mu^*(T \cap E_1 \cup E_3^C \cup T \cap E_2 \cap E_3^C) \\ &= \mu^*(T \cap E_3) + \mu^*(T \cap E_1 \cup T \cap E_2),\end{aligned}$$

because $E_1 \subseteq E_3^C$ and $E_2 \subseteq E_3^C$ since all the E_n are disjoint. Then, we can apply the first step to conclude

$$\mu^*(T \cap (E_1 \cup E_2 \cup E_3)) = \mu^*(T \cap E_3) + \mu^*(T \cap E_2) + \mu^*(T \cap E_1).$$

We have therefore shown

$$\mu^*(T \cap (\cup_{i=1}^3 E_i)) = \sum_{i=1}^{3} \mu^*(T \cap E_i).$$

It is now clear, we can continue this argument by induction to show

$$\mu^*(T \cap (\cup_{i=1}^n E_i)) = \sum_{i=1}^{n} \mu^*(T \cap E_i). \qquad (\boldsymbol{a})$$

for any positive integer n. Further, since $\mathcal{M}$ is an algebra, induction also shows $\cup_{i-1}^n E_i$ is in $\mathcal{M}$ for any such n. It then follows that for any T in X,

$$\mu^*(T) = \mu^*(T \cap (\cup_{i=1}^n E_i)) + \mu^*(T \cap (\cup_{i=1}^n E_i)^C).$$

*Using Equation **a**, we then have*

$$\mu^*(T) = \sum_{i=1}^{n} \mu^*(T \cap E_i) + \mu^*(T \cap (\cup_{i=1}^n E_i)^C). \qquad (\boldsymbol{b})$$

Next, note for all n

$$T \cap (\cup_{i=1}^n E_i)^C \supseteq T \cap (\cup_{i=1}^\infty E_i)^C,$$

and hence

$$\mu^*\Big(T \cap (\cup_{i=1}^\infty E_i)^C\Big) \leq \mu^*\Big(T \cap (\cup_{i=1}^n E_i)^C\Big).$$

*Using this in Equation **b**, we find*

$$\mu^*(T) \geq \sum_{i=1}^{n} \mu^*(T \cap E_i) + \mu^*(T \cap (\cup_{i=1}^i nfty\, E_i)^C). \qquad (\boldsymbol{c})$$

Since this holds for all n, letting $n \to \infty$, we obtain

$$\mu^*(T) \geq \sum_{i=1}^{\infty} \mu^*(T \cap E_i) + \mu^*(T \cap (\cup_{i=1}^i nfty\, E_i)^C). \qquad (\boldsymbol{d})$$

Finally, since

$$\bigcup_{i=1}^{\infty}(T \cap E_i) = T \bigcap (\cup_{i=1}^\infty E_i),$$

by the countable subadditivity of μ^**, it follows that*

$$\mu^*\Big(T \bigcap (\cup_{i=1}^{\infty} E_i)\Big) \quad = \quad \mu^*\Big(\bigcup_{i=1}^{\infty}(T \cap E_i)\Big) \leq \sum_{i=1}^{\infty} \mu^*\Big(T \cap E_i)\Big).$$

Using this in Equation **c***, we have*

$$\mu^*(T) \quad \geq \quad \mu^*\Big(T \bigcap (\cup_{i=1}^{\infty} E_i)\Big) + \mu^*\Big(T \cap (\cup_{i=1}^{i} nfty\, E_i)^C\Big). \qquad (\boldsymbol{e})$$

Since this holds for all subsets T*, this tells us* $\cup_n E_n$ *is in* $\mathcal{M}$*. This proves that* $\mathcal{M}$ *is a* σ *- algebra.*

However, with all this work already done, we can also derive a very nice result which will help us later. Countable subadditivity of μ^* *gives us*

$$\mu^*(T) \quad \leq \quad \mu^*\Big(T \bigcap (\cup_{i=1}^{\infty} E_i)\Big) + \mu^*\Big(T \cap (\cup_{i=1}^{i} nfty\, E_i)^C\Big).$$

Hence, using countable subadditivity again,

$$\mu^*(T) \quad \leq \quad \sum_{i=1}^{\infty} \mu^*\Big(T \cap E_i)\Big) + \mu^*\Big(T \cap (\cup_{i=1}^{i} nfty\, E_i)^C\Big). \qquad (\boldsymbol{f})$$

Combining Equation **d** *and Equation* ***f****, we find*

$$\mu^*(T) \quad = \quad \sum_{i=1}^{\infty} \mu^*\Big(T \cap E_i)\Big) + \mu^*\Big(T \cap (\cup_{i=1}^{i} nfty\, E_i)^C\Big).$$

Thus, letting "T*" be "*$T \cap (\cup_n E_n)$*", we find*

$$\mu^*(T \cap \cup_{i=1}^{\infty} E_i) \quad = \quad \sum_{i=1}^{\infty} \mu^*\Big(T \cap E_i)\Big). \qquad (\boldsymbol{g})$$

■

Theorem 12.1.3. The Measure Induced By An Outer Measure

Let X *be a non empty set,* μ^* *an outer measure on* X *and* $\mathcal{M}$ *be the collection of* μ^* *measurable subsets of* X*. Then,* $\mathcal{M}$ *is a* σ *- algebra and* μ^* *restricted to* $\mathcal{M}$ *is a measure we will denote by* μ*.*

Proof. *Recall that* $\mathcal{M}$ *is a* σ *- algebra if*

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

(ii) If $A \in \mathcal{M}$*, so is* A^C.

(iii) If $\{A_n\}_{n=1}^{\infty} \in \mathcal{M}$*, then* $\cup_{n=1}^{\infty} A_n \in \mathcal{M}$.

Since we know $\mathcal{M}$ *is an algebra of sets, all that remains is to show it is closed under countable unions. We have already shown all the properties of a* σ *- algebra except closure under arbitrary countable unions.*

The previous theorem, however, does give us closure under countable disjoint unions. So, let (A_n) be a countable collection of sets in $\mathcal{M}$. Letting

$$\begin{aligned} E_1 &= A_1 \\ E_2 &= A_2 \setminus A_1 \\ \vdots &= \vdots \\ E_n &= A_n \setminus \left(\cup_{i=1}^{n-1} A_i \right) \\ \vdots &= \vdots, \end{aligned}$$

we see each E_n is in $\mathcal{M}$ by Theorem 12.1.1. Further, they are pairwise disjoint and so by Theorem 12.1.2, we can conclude $\cup_n E_n$ is in $\mathcal{M}$. But it is easy to see that $\cup_n E_n = \cup_n A_n$. Thus, $\mathcal{M}$ is a σ - algebra.

To show μ^* restricted to $\mathcal{M}$, μ, is a measure, we must show

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

(ii): $\mu(E) \geq 0$, *for all* $E \in \mathcal{S}$,

(iii): μ *is countably additive on* $\mathcal{S}$*; i.e. if* $(E_n) \subseteq \mathcal{S}$ *is a countable collection of disjoint sets, then* $\mu(\cup_n E_n) = \sum_n \mu(E_n)$.

Since $\mu^*(\emptyset) = 0$, condition (i) follows immediately. Also, we know $\mu^*(E) \geq 0$ for all subsets E, and so condition (ii) is valid. It remains to show countable additivity. Let (B_n) be a countable disjoint family in $\mathcal{M}$. We can apply Equation ***g*** *to conclude, using $T = X$, that*

$$\mu^*(\cap \cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu^*\Big(\cap B_i\Big).$$

Finally, since $\mu^* = \mu$ on these sets, we have shown μ is countably additive and so is a measure. ■

It is also true that the measure constructed from an outer measure in this fashion is a complete measure.

Theorem 12.1.4. The Measure Induced By An Outer Measure Is Complete

If E is a subset of X satisfying $\mu^*(E) = 0$, then $E \in \mathcal{M}$. Also, if $F \subseteq E$, then $F \in \mathcal{M}$ as well, with $\mu^*(F) = 0$. Note, this tells us that if $\mu(E) = 0$, then subsets of E are also in $\mathcal{M}$ with $\mu(F) = 0$; i.e., μ is a complete measure.

Proof. *We know $\mu^*(T \cap E) \leq \mu^*(E)$ for all T; hence, $\mu^*(T \cap E) = 0$ here. Thus, for any T,*

$$\begin{aligned} \mu^*(T \cap E) + \mu^*(T \cap E^C) & \\ &= \mu^*(T \cap E^C) \leq \mu^*(T). \end{aligned}$$

This tells us E satisfies the Caratheodory condition and so is in $\mathcal{M}$. Thus, we have $\mu(E) = 0$. Now, let $F \subseteq E$. Then, $\mu^*(F) = 0$ also; hence, by the argument above, we can conclude $F \in \mathcal{M}$ with $\mu(F) = 0$. ■

12.2 Measures From Metric Outer Measures

Definition 12.2.1. Metric Outer Measure
Let (X, d) be a non empty metric space and for two subsets A and B of X, define the distance between A and B by

$$D(A, B) \quad = \quad \inf \{ \, d(a, b) \, | \, a \, \in A, \;\; b \, \in B \, \}.$$

If μ^* is an outer measure on X which satisfies

$$\mu^*(A \, \cup \, B) \quad = \quad \mu^*(A) \, + \, \mu^*(B)$$

whenever $D(A, B) > 0$, we say μ^* is metric outer measure.

The σ algebra of open subsets of X is called the Borel σ algebra $\mathcal{B}$. We can use the construction process in Section 12.1 to construct a σ algebra of subsets, $\mathcal{M}$, which satisfy the Caratheodory condition for this metric outer measure μ^*. This gives us a measure on $\mathcal{M}$. We would like to be able to say that open sets in the metric spaced are μ^* measurable. Thus, we want to prove $\mathcal{B} \subseteq \mathcal{M}$. This is what we do in the next theorem. It is becoming a bit cumbersome to keep saying μ^* measurable for the sets in $\mathcal{M}$. We will make the following convention for later use: a set in $\mathcal{M}$ will be called **OMI** measurable, where **OMI** stands for outer measure induced.

Theorem 12.2.1. Open Sets in a Metric Space Are OMI Measurable
Let (X, d) be a non empty metric space and μ^* a metric outer measure on X. Then open sets are OMI measurable.

Proof. *let E be open in X. To show E is μ^* measurable we must show*

$$\mu^*(T) \quad \geq \quad \mu^*(T \cap E) \, + \, \mu^*(T \cap E^C)$$

for all subsets T in X. Since this is true for all subsets with $\mu^*(T) = \infty$, it suffices to prove the inequality is valid for all subsets with $\mu^*(T)$ finite. Also, we already know $\emptyset$ and X are μ^* measurable, so we can further restrict our attention to nonempty strict subsets E of X. We will prove this in a series of steps:

Step (i): *Let E_n be defined for each positive integer n by*

$$E_n \quad = \quad \{ \, x \, | \, D(x, E^C) \, > \, \frac{1}{n} \}.$$

It is clear $E_n \subseteq E$ and that $E_n \subseteq E_{n+1}$.

Note, if $y \in E_n$ and $x \in E^c$, we have $d(y, x) > 1/n$ and so

$$\inf_{y \in E_n, \, x \in E^c} d(y, x) \quad \geq \quad \frac{1}{n}$$

and so $D(E_n, E^C) \geq 1/n$. This immediately tells us

$$D(T \, \cap \, E_n, T \cap \, E^C) \geq 1/n$$

also for all T.

Since μ^* is a metric outer measure, we then have

$$\mu^*\Big((T \cap E_n) \cup (T \cap E^C)\Big) \;=\; \mu^*(T \cap E_n) + \mu^*(T \cap E^C).$$

However, we also know E_n is a subset of E and so

$$(T \cap E_n) \cup (T \cap E^C) \;\subseteq\; (T \cap E) \cup (T \cap E^C) = T.$$

We conclude then

$$\mu^*\Big((T \cap E_n) \cup (T \cap E^C)\Big) \;\leq\; \mu^*(T).$$

Hence, for all T, we have

$$\mu^*\Big(T \cap E_n)\Big) + \mu^*\Big(T \cap E^C)\Big) \;\leq\; \mu^*(T). \tag{$*$}$$

Step (ii): *If $\lim_n \mu^*(T \cap E_n) = \mu^*(T)$, then letting n go to infinity in Equation $*$, we would find*

$$\mu^*\Big(T \cap E)\Big) + \mu^*\Big(T \cap E^C)\Big) \;\leq\; \mu^*(T).$$

This means E satisfies the Caratheodory condition and so is μ^* measurable.

To show this limit acts in this way, we will construct a new sequence of sets (W_n) that are disjoint from one another with $E = cup_n W_n$ so that the new sets W_n have useful properties. Since E is open, every point p in E is an interior point. Thus, there is a positive r so that $B(p;r) \subseteq E$. So, if $z \in E^C$, we must have and $d(p,z) \geq r$. It follows that $D(p, E^C) \geq r > r/2$. We therefore know that $p \in E_n$ for some n. Since our choice of p is arbitrary, we have shown

$$E \;\subseteq\; \cup_n E_n.$$

It was already clear that $\cup_n E_n \subseteq E$; we conclude $E = \cup_n E_n$. We then define the needed disjoint collection (W_n) as follows

$$\begin{array}{rcl} W_1 & = & E_1 \\ W_2 & = & E_2 \setminus E_1 \\ W_2 & = & E_3 \setminus E_2 \\ \vdots & \vdots & \vdots \\ W_n & = & E_n \setminus E_{n-1} \end{array}$$

(It helps to draw a picture here for yourself in terms of the annuli $E_n \setminus E_{n-1}$. We can see that for any n, we can write

$$T \cap E \quad = \quad (T \cap E_n) \bigcup \cup_{k=n+1}^{\infty} (T \cap W_k)$$

as the terms $T \cap W_k$ give the contributions of each annuli or strip outside of the core E_n. Hence,

$$\mu^*(T \cap E) \quad \leq \quad \mu^*(T \cap E_n) + \sum_{k=n+1}^{\infty} (T \cap W_k) \qquad (**)$$

because μ^* is subadditive. At this point, the series sum $\sum_{k=n+1}^{\infty} (T \cap W_k)$could be ∞; we haven't determined if it is finite yet.

For any $k > 1$, if $x \in W_k$, then $x \in E_k \setminus E_{k-1}$ and so

$$\frac{1}{k} \quad \leq \quad D(x,\, E^C) \;\leq\; \frac{1}{k-1}.$$

Next, if $x \in W_k$ and $y \in W_{k+p}$ for any $p \geq 2$, we can use the triangle inequality with an arbitrary $z \in E^C$ to conclude

$$d(x,z) \quad \leq \quad d(x,y) + d(y,z).$$

But, this says

$$\begin{aligned} d(x,y) \quad &\geq \quad d(x,z) - d(y,z) \\ &\geq \quad D(x, E^C) - d(y,z) \;>\; \frac{1}{k} - d(y,z). \end{aligned}$$

We have shown the fundamental inequality

$$d(x,y) \quad > \quad \frac{1}{k} - d(y,z),\, \forall\, x \in W_k, \forall\, y \in W_{k+p} \qquad (\boldsymbol{\alpha})$$

holds for $p \geq 2$. The definition of the set E_{k+p} then implies for these p,

$$\frac{1}{k+p} \quad < \quad D(y, E^C) \;\leq\; \frac{1}{k+p-1}. \qquad (\boldsymbol{\beta})$$

Now consider how $D(y, E^C)$ is defined. Since this is an infimum, by the Infimum Tolerance Lemma, given a positive ϵ, there is a $z_\epsilon \in E^C$ so that

$$D(y, E^C) \quad \leq \quad d(y, z_\epsilon) \;<\; D(y, E^C) + \epsilon.$$

Hence, using Equation $\boldsymbol{\beta}$*, we have*

$$\begin{aligned} -d(y, z_\epsilon) &> -D(y, E^C) - \epsilon \\ &> -\frac{1}{k+p-1} - \epsilon. \end{aligned}$$

Also, using Equation $\boldsymbol{\alpha}$*, we find*

$$\begin{aligned} d(x, y) &> \frac{1}{k} - d(y, z_\epsilon) \\ &> \frac{1}{k} - -\frac{1}{k+p-1} - \epsilon \\ &= \frac{p-1}{k(k+p-1)} - \epsilon. \end{aligned}$$

Since $\epsilon > 0$ *is arbitrary, we conclude*

$$d(x, y) \geq \frac{p-1}{k(k+p-1)} > 0$$

for all $x \in W_k$ *and* $y \in W_{k+p}$ *with* $p \geq 2$*. Hence,*

$$D(W_k,\, W_{k+p}) \geq \frac{p-1}{k(k+p-1)} > 0$$

It follows that

$$D(W_1,\, W_3) > 0$$

and, in general, we find this is true for the successive odd integers

$$D(W_{2k+1},\, W_{2k+3}) > 0.$$

Since μ^* *is a metric outer measure, this allows us to say*

$$\begin{aligned} \sum_{k=0}^{n} \mu^*(T \cap W_{2k+1}) &= \mu^*\Big(\cup_{k=0}^{n} T \cap W_{2k+1}\Big) \\ &\leq \mu^*\Big(\cup_{k=0}^{\infty} T \cap W_{2k+1}\Big) \leq \mu^*(T). \end{aligned}$$

A similar argument shows that successive even integers satisfy

$$D(W_{2k},\, W_{2k+2}) > 0.$$

Again, as μ^* *is a metric outer measure, this allows us to say*

$$\sum_{k=0}^{n} \mu^*(T \cap W_{2k}) = \mu^*\Big(\cup_{k=0}^{n} T \cap W_{2k}\Big).$$

Therefore, we have

$$\begin{aligned}\sum_{k=0}^{n} \mu^*(T \cap W_{2k}) &\leq \mu^*\Big(\cup_{k=0}^{\infty} T \cap W_{2k}\Big) \\ &\leq \mu^*(T).\end{aligned}$$

We conclude

$$\begin{aligned}\sum_{k=0}^{n} \mu^*(T \cap W_k) &= \sum_{k \text{ even}} \mu^*(T \cap W_k) + \sum_{k \text{ odd}} \mu^*(T \cap W_k) \\ &\leq 2\,\mu^*(T)\end{aligned}$$

for all n. This implies the sum $\sum_k \mu^*(T \cap W_k)$ converges to a finite number.

Since the series converges, we now know given $\epsilon > 0$, there is an N so that

$$\sum_{k=n}^{\infty} \mu^*(T \cap W_k) < \epsilon,$$

for all $n > N$. Now go back to Equation $**$. We have for any $n > N$,

$$\mu^*(T \cap E) \leq \mu^*(T \cap E_n) + \epsilon.$$

This tells us

$$\mu^*(T \cap E) \leq \mu^*(T \cap E_n), \forall\, n > N,$$

or $\mu^*(T \cap E_n) \to \mu^*(T \cap E)$. By our earlier remark, this completes the proof. ■

We can even prove more.

Theorem 12.2.2. Open Sets In A Metric Space Are μ^* Measurable If and Only If $\mu^a st$ Is A Metric Outer Measure

Let X be a non empty metric space. Then Open sets are μ^* measurable if and only if μ^* is a metric outer measure.

Proof. *If we assume μ^* is a metric outer measure, then opens sets are μ^* measurable by Theorem 12.2.1.*

On the other hand if we know that all the open sets of μ^* measurable, this implies all Borel sets are μ^* measurable as well. Let A and B be any two sets with $D(A, B) = r > 0$. For each $x \in A$, let

$$G(x) = \{\, u \mid d(x, u) < r/2 \,\}$$

and

$$G = \bigcup_{x \in A} G(x).$$

Then G is open, $A \subseteq G$ and $G \cap B = \emptyset$. Since G is measurable, it satisfies the Caratheodory condition using test set $T = A \cup B$; thus,

$$\mu^*\Big(A \cup B\Big) \;=\; \mu^*\Big((A \cup B\Big) \cap G) + \mu^*\Big((A \cup B\Big) \cap G^C).$$

But $(A \cup B\Big) \cap G$ is simplified to A because $A \subseteq B$ and B is disjoint from G. Further since A is disjoint from G^C and $B \subseteq G^C$, we have $(A \cup B\Big) \cap G^C = B$. We conclude

$$\mu^*\Big(A \cup B\Big) \;=\; \mu^*\Big(A) + \mu^*\Big(B).$$

This shows μ^* is a metric outer measure. ∎

12.3 Constructing Outer Measures

We still have to find ways to construct outer measures. We want the resulting OMI measure we induce have certain properties useful to us. Let's discuss how to do this now.

Definition 12.3.1. Premeasures and Covering Families

Let X be a nonempty set. Let $\mathscr{T}$ be a family of subsets of X that contains the empty set. This family is called a covering family for X. Let τ be a mapping on $\mathscr{T}$ so that $\tau(\emptyset) = 0$. The mapping τ is called a premeasure.

It is hard to believe, but even with virtually no restrictions on τ and $\mathscr{T}$, we can build an outer measure.

Theorem 12.3.1. Constructing Outer Measures Via Premeasures

Let X be a nonempty set. Let $\mathscr{T}$ be a covering family of subsets of X and $\tau : \mathscr{T} \to [0, \infty]$ be a premeasure. For any A in X, define

$$\mu^*(A) \;=\; \inf \{ \sum_n \tau(T_n) \,|\, T_n \in \mathscr{T},\, A \subseteq \cup_n T_n \}$$

where the sequence of sets (T_n) from $\mathscr{T}$ is finite or countably infinite. Such a sequence is called a covering family. In the case where there are no sets from $\mathscr{T}$ that cover A, we define the infimum over the resulting empty set to be ∞. Then μ^* is an outer measure on X.

Proof. *To verify the mapping μ^* is an outer measure on X, we must show*

(i): $\mu^*(\emptyset) = 0$.

(ii): If A and B are subsets of X with $A \subseteq B$, then $\mu^*(A) \leq \mu^*(B)$.

(iii): If (A_n) is a sequence of disjoint subsets of X, then $\mu^*(\cup_n A_n) \leq \sum_n \mu^*(A_n)$.

It is straightforward to see condition (i) and (ii) are true. It suffices to prove condition (iii) is valid. Let (A_n) be a countable collection, finite or infinite, of subsets of X. If there is an index n with $\tau(A_n)$ infinite, then since $\mu^*(\cup_n A_n) \leq \infty$ anyway, it is clear

$$\mu^*(\cup_{i=1}^{\infty} A_i) \quad \leq \quad \sum_{i=1}^{\infty} \mu^*(A_i) \;=\; \infty.$$

On the other hand, if $\mu^*(A_n)$ is finite for all n, given any $\epsilon > 0$, we can use the Infimum Tolerance Lemma to find a sequence of families $(T_{n\,k})$ in $\mathscr{T}$ so that

$$\sum_{k=1}^{\infty} \tau(T_{n\,k}) \quad < \quad \mu^*(A_n) \,+\, \frac{\epsilon}{2^n}.$$

We also know that

$$\bigcup_{n=1}^{\infty} A_n \quad \subseteq \quad \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} T_{n\,k}.$$

Hence, the collection $\cup_n \cup_k T_{n\,k}$ is a covering family for $\cup_n A_n)$ and so by the definition of μ^*, we have

$$\begin{aligned} \mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \quad &\leq \quad \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu^*\Big(T_{n\,k}\Big) \\ &\leq \quad \sum_{n=1}^{\infty} \{\mu^*(A_n) \,+\, \frac{\epsilon}{2^n}\} \\ &\leq \quad \sum_{n=1}^{\infty} \mu^*(A_n) \,+\, \epsilon. \end{aligned}$$

Since ϵ is arbitrary, we see μ^* is countable subadditive and so is an outer measure. ∎

There is so little known about τ and $\mathscr{T}$, that it is not clear at all that

(i): $\mathscr{T} \subseteq \mathcal{M}$, where $\mathcal{M}$ is the σ - algebra of sets that satisfy the Caratheodory condition for the outer measure μ^* generated by τ. If this is true, we will call $\mathcal{M}$ an OMI-F σ - algebra, where the "F" denotes the fact that the covering family is an algebra.

(ii): If $A \in \mathscr{T}$, then $\tau(A) = \mu(A)$ where μ is the measure obtained by restricting μ^* to $\mathcal{M}$. If this is true, we will call the constructed σ - algebra, an OMI-FE σ - algebra, where the "E" indicates the fact the μ restricted to $\mathscr{T}$ recovers τ.

If τ represents some primitive notion of size of special sets, like length of intervals on the real line, we normally want both condition (i) and (ii) above to be valid. We can obtain these results if we add a few more properties to τ and $\mathscr{T}$. First, $\mathscr{T}$ needs to be an algebra (which we have already defined) and τ needs to be additive on the algebra.

Definition 12.3.2. Additive Set Function

Let $\mathcal{A}$ of subsets of the set X be an algebra. Let ν be an extended real valued function defined on $\mathcal{A}$ which satisfies

(i): $\nu(\emptyset) = 0$.

(ii): If A and B in $\mathcal{A}$ are disjoint, then $\nu(A \cup B) = \nu(A) + \nu(B)$.

Then ν is called an additive set function on $\mathcal{A}$.

We also need a property of outer measures called regularity.

Definition 12.3.3. Regular Outer Measures

Let X be a nonempty set, μ^* be an outer measure on X and $\mathcal{M}$ be the set of all μ^* measurable sets of X. The outer measure μ^* is called regular if for all E in X there is a μ^* measurable $F \in \mathcal{M}$ so that $E \subseteq F$ with $\mu^*(E) = \mu(F)$, where μ is the measure induced by μ^* on $\mathcal{M}$. The set F is often called a measurable cover for E.

We begin with a technical lemma.

Lemma 12.3.2. Condition For Outer Measure To Be Regular

Let X be a nonempty set, $\mathscr{T}$ a covering family and τ a premeasure. Then if the σ - algebra, $\mathcal{M}$, generated by τ using $\mathscr{T}$ contains $\mathscr{T}$, μ^* is regular.

Proof. *Let A be a subset in X. We need to show there is measurable set B containing A so that $\mu^*(A) = \mu(B)$. If the $mu^*(A) = \infty$, then we can choose X as the needed set. Otherwise, we have $\mu^*(A)$ is finite. Applying the Infimum Tolerance Lemma, for each m, there is a family of sets (E_n^m) so that $A \subseteq \cup_n E_n^m)$ and*

$$\sum_n \tau(E_n^m) \quad < \quad \mu^*(A) + \frac{1}{m}.$$

Let

$$\begin{aligned} E_m &= \bigcup_n E_n^m \\ H &= \bigcap_m E_m; \end{aligned}$$

these sets are measurable by assumption. Also, $A \subseteq H$ and $H \subseteq E_m$. Hence, $\mu^*(A) \leq \mu(H)$. We now show the reverse inequality. For each m, we have

$$\begin{aligned} \mu^*(E_m) &\leq \sum_n \mu^*(E_n^m) \leq \sum_n \tau(E_n^m) \\ &\leq \mu^*(A) + \frac{1}{m}. \end{aligned}$$

Further, since $H \subseteq E_m$ for each m, we find

$$\mu(H) \quad \leq \quad \mu^*(E_m) \leq \mu^*(A) + \frac{1}{m}.$$

This is true for all m; hence, it follows that $\mu(H) \leq \mu^*(A)$. Combining inequalities, we have $\mu(H) = \mu^*(A)$ and so H is a measurable cover. Thus, μ^* is regular. ■

Theorem 12.3.3. Conditions For OMI-F Measures
Let X be a nonempty set, $\mathscr{T}$ a covering family which is an algebra and τ an additive set function on $\mathscr{T}$. Then the σ - algebra, $\mathcal{M}$, generated by τ using $\mathscr{T}$ contains $\mathscr{T}$ and μ^* is regular.

Proof. *By Lemma 12.3.2, it is enough to show each member of $\mathscr{T}$ is measurable. So, let A be in $\mathscr{T}$. As usual, it suffices to show that*

$$\mu^*(T) \;\geq\; \mu^*(T \cap A) + \mu^*(T \cap A^C)$$

for all sets T of finite outer measure. This will show A satisfies the Caratheodory condition and hence, is measurable. Let $\epsilon > 0$ be given. By the Infimum Tolerance Lemma, there is a family (A_n) from $\mathscr{T}$ so that $T \subseteq \cup_n A_n$ and

$$\sum_n \tau(A_n) \;<\; \mu^*(T) + \epsilon.$$

since τ is additive on $\mathscr{T}$, we know

$$\tau(A_n) \;=\; \tau(A \cap A_n) + \tau(A^C \cap A_n).$$

Also, we have

$$A \bigcap T \;\subseteq\; \bigcup_n (A \cap A_n), \quad \textit{and} \quad A^C \bigcap T \subseteq \bigcup_n (A^C \cap A_n).$$

Hence,

$$\mu^*(A \cap T) \;\; leq \;\; \sum_n \mu^*(A \cap A_n), \;\; \mu^*(A^C \cap T) \leq \sum_n \mu^*(A^C \cap A_n). \qquad (\boldsymbol{\alpha})$$

$$\begin{aligned} \mu^*(T) + \epsilon \;&>\; \sum_n \tau(A_n) = \sum_n \tau(A_n \cap A) + \sum_n \tau(A_n \cap A^C) \\ &\geq\; \sum_n \mu^*(A_n \cap A) + \sum_n \mu^*(A_n \cap A^C) \\ &\geq\; \mu^*(A \cap T) + \mu^*(A^C \cap T), \end{aligned}$$

by Equation $\boldsymbol{\alpha}$. Thus, A satisfies the Caratheodory condition and is measurable. ■

In order for condition (ii) to hold, we need to add one more additional property to τ: it needs to be a pseudo-measure.

Definition 12.3.4. Pseudo-Measure
Let the mapping $\tau : \mathcal{A} \rightarrow [0, \infty]$ be additive on the algebra $\mathcal{A}$. Assume whenever (A_i) is a countable collection of disjoint sets in $\mathcal{A}$ whose union is also in $\mathcal{A}$ (note this is not always true because $\mathcal{A}$ is not a σ - algebra), then it is true that

$$\tau(\cup_i A_i) \quad = \quad \sum_i \tau(A_i).$$

Such a mapping τ is called a pseudo-measure on $\mathcal{A}$.

Theorem 12.3.4. Conditions For OMI-FE Measures
Let X be a nonempty set, $\mathscr{T}$ a covering family which is an algebra and τ an additive set function on $\mathscr{T}$ which is a pseudo-measure. Then the σ - algebra, $\mathcal{M}$, generated by τ using $\mathscr{T}$ contains $\mathscr{T}$, μ^* is regular and $\mu(T) = \tau(T)$ for all T in $\mathcal{A}$.

Proof. *see Bruckner.* ■

Comment 12.3.1. *The results above tell us that we can construct measures satisfying condition (i) and (ii) as long as the premeasure is a pseudo-measure and the covering family is an algebra. This means the covering family must be closed under complementation. Hence, if we a covering family such as the collection of all open intervals (which we do when we construct Lebesgue measure later) these theorems do not apply.*

12.4 Worked Out Problems

Let's work out a specific examples of this process to help the ideas sink in. Note the covering families here to not simply contain open intervals!

Example 12.4.1. *Let $\mathcal{U}$ be the family of subsets of $\Re$ of the form $(a, b]$, $(-\infty, b]$, (a, ∞) and $(-\infty, \infty)$. It is easy to show that $\mathcal{F}$, the collection of all finite unions of sets from $\mathcal{U}$ is an algebra of subsets of $\Re$. Let τ be the usual length of an interval. and extend τ to $\mathcal{F}$ additively. This extended τ is a premeasure on $\mathcal{F}$. τ can then be used to define an outer measure as usual $\mu^*(\tau)$. There is then an associated σ - algebra of μ^*_τ measurable sets of $\Re$, $\mathcal{M}_\tau$, and μ^*_τ restricted to $\mathcal{M}_\tau$ is a measure is a measure, μ_τ.*
We will now prove $\mathcal{F}$ is contained in $\mathcal{M}_\tau$. Let's consider the set I from $\mathcal{U}$. Let T be any subset of $\Re$ and let $\epsilon > 0$ be given. Then there is a cover (A_n) of sets from the algebra $\mathcal{F}$ so that

$$\sum_n \tau(A_n) \quad \leq \quad \mu^*_\tau(T) \, + \, \epsilon.$$

Now $I \cap T \subseteq \cup_n (A_n \cap I)$ and $I^C \cap T \subseteq \cup_n (A_n \cap I^C)$. So because $\mathcal{F}$ is an algebra, this means $(A_n \cap I)$ covers $I \cap T$ and $(A_n \cap I^C)$ covers $I^C \cap T$. Hence,

$$\begin{aligned} \mu^*_\tau(T \cap I) \quad &\leq \quad \sum_n \tau(A_n \cap I), \\ \mu^*_\tau(T \cap I^C) \quad &\leq \quad \sum_n \tau(A_n \cap I^C). \end{aligned}$$

Combining, we see

$$\mu_\tau^*(T \cap I) + \mu_\tau^*(T \cap I^C) \leq \sum_n \Big(\tau(A_n \cap I) + \tau(A_n \cap I^C) \Big).$$

But τ is additive on $\mathcal{F}$, and hence

$$\sum_n \Big(\tau(A_n \cap I) + \tau(A_n \cap I^C) \Big) = \sum_n \tau(A_n).$$

Thus,

$$\mu_\tau^*(T \cap I) + \mu_\tau^*(T \cap I^C) \leq \mu_\tau^*(T) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have shown I satisfies the Caratheodory condition. This shows that I is OMI measurable and so $\mathcal{F} \subseteq \mathcal{M}_\tau$.

Example 12.4.2. *Let $\mathcal{U}$ be the family of subsets of $\Re$ of the form $(a, b]$, $(-\infty, b]$, (a, ∞) and $(-\infty, \infty)$ and the empty set. It is easy to show that $\mathcal{F}$, the collection of all finite unions of sets from $\mathcal{U}$ is an algebra of subsets of $\Re$. Let g be the monotone increasing function on $\Re$ defined by $g(x) = x^2$. Note g is right continuous which means*

$$\begin{aligned} \lim_{h \to 0^+} g(x+h) \ \ &\textit{exists}, \ \forall\, x, \\ \lim_{x \to -\infty} g(x) \ \ &\textit{exists}, \\ \lim_{x \to \infty} g(x) \ \ &\textit{exists}. \end{aligned}$$

where the last two limits are $-\infty$ and ∞ respectively. Define the mapping τ_g on $\mathcal{U}$ by

$$\begin{aligned} \tau_g\Big((a, b] \Big) &= g(b) - g(a), \\ \tau_g\Big((-\infty, b) \Big) &= g(b) - \lim_{x \to -\infty} g(x), \\ \tau_g\Big((a, \infty) \Big) &= \lim_{x \to \infty} g(x) - g(a), \\ \tau_g\Big((-\infty, \infty) \Big) &= \lim_{x \to \infty} g(x) - \lim_{x \to -\infty} g(x). \end{aligned}$$

Extend τ_g to $\mathcal{F}$ additively as usual. This extended τ_g is a premeasure on $\mathcal{F}$. τ_g can then be used to define an outer measure as usual $\mu^*(g)$. There is then an associated σ - algebra of μ_g^* measurable sets of $\Re$, $\mathcal{M}_g$, and μ_g^* restricted to $\mathcal{M}_g$ is a measure, μ_g.

We will now prove $\mathcal{F}$ is contained in $\mathcal{M}_g$. Let's consider the set I from $\mathcal{U}$. Let T be any subset of $\Re$ and let $\epsilon > 0$ be given. Then there is a cover (A_n) of sets from the algebra $\mathcal{F}$ so that

$$\sum_n \tau_g(A_n) \leq \mu_g^*(T) + \epsilon.$$

Now $I \cap T \subseteq \cup_n (A_n \cap I)$ *and* $I^C \cap T \subseteq \cup_n (A_n \cap I^C)$. *So*

$$\begin{aligned} \mu_g^*(T \cap I) &\leq \sum_n \tau_g(A_n \cap I), \\ \mu_g^*(T \cap I^C) &\leq \sum_n \tau_g(A_n \cap I^C). \end{aligned}$$

Combining, we see

$$\mu_g^*(T \cap I) + \mu_g^*(T \cap I^C) \leq \sum_n \Big(\tau_g(A_n \cap I) + \tau_g(A_n \cap I^C) \Big).$$

But τ_g *is additive on* $\mathcal{F}$*, and hence*

$$\sum_n \Big(\tau_g(A_n \cap I) + \tau_g(A_n \cap I^C) \Big) = \sum_n \tau_g(A_n).$$

Thus,

$$\mu_g^*(T \cap I) + \mu_g^*(T \cap I^C) \leq \mu_g^*(T) + \epsilon.$$

Since $\epsilon > 0$ *is arbitrary, we have shown* I *satisfies the Caratheodory condition. This shows that* I *is OMI measurable and so* $\mathcal{F} \subseteq \mathcal{M}_g$.

12.5 Homework

Exercise 12.5.1. *Let* $X = (0, 1]$. *Let* $\mathcal{A}$ *consist of the empty set and all finite unions of half- open intervals of the form* $(a, b]$ *from* X. *Prove* $\mathcal{A}$ *is an algebra of sets of* $(0, 1]$.

Exercise 12.5.2. *Let* $\mathcal{A}$ *be the algebra of subsets of* $(0, 1]$ *given in Exercise 12.5.1. Let* f *be an arbitrary function on* $[0, 1]$. *Define* ν_f *on* $\mathcal{A}$ *by*

$$\nu_f\Big((a, b]\Big) = f(b) - f(a).$$

Extend ν_f *to be additive on finite disjoint intervals as follows: if* $(A_i) = (a_i, b_i])$ *is a finite collection of disjoint intervals of* $(0, 1]$, *we define*

$$\nu_f\Big(\cup_{i=1}^n (a_i, b_i]\Big) = \sum_{i=1}^n f(b_i) - f(a_i).$$

1. *Prove that* ν_f *is additive on* $\mathcal{A}$.

   Hint. *It is enough to show that the value of* $\nu_f(A)$ *is independent of the way in which we write* A *as a finite disjoint union.*

2. *Prove* ν_f *is non negative if and only if* f *is non decreasing.*

Exercise 12.5.3. *If* λ *is an additive set function on an algebra of subsets* $\mathcal{A}$*, prove that* λ *can not take on both the value* ∞ *and* $-\infty$.

Hint. *If there is a set A in the algebra with $\lambda(A) = \infty$ and there is a set B in the algebra with $\lambda(B) = -\infty$, then we can find disjoint sets A' and B' in $\mathcal{A}$ so that $\lambda(A') = \infty$ and $\lambda(B') = -\infty$. But this is not permitted as the value of $\lambda(A' \cup B')$ must be a well - defined extended real value not the undefined value $\infty - \infty$.*

Exercise 12.5.4. *Let $\mathscr{T}$ be a covering family for a nonempty set X. Let τ be a non negative, possibly infinite valued premeasure. For any A in X, define*

$$\mu^*(A) \quad = \quad \inf \{ \sum_n \tau(T_n) \,|\, T_n \in \mathscr{T}, \, A \subseteq \cup_n T_n \}$$

where the sequence of sets (T_n) from $\mathscr{T}$ is finite or countably infinite. In the case where there are no sets from $\mathscr{T}$ that cover A, we define the infimum over the resulting empty set to be ∞.

Prove μ^* is an outer measure on X.

Exercise 12.5.5. *Let $X = \{1. 2, 3\}$ and $\mathscr{T}$ consist of $\emptyset$, X and all doubleton subsets $\{x, y\}$ of X. Let τ satisfy*

(i): $\tau(\emptyset) = 0$.

(ii): $\tau\Big(\{x, y\}\Big) = 1$ *for all* $x \neq y$ *in* X.

(iii): $\tau(X) = 2$.

(a): *Prove the method of Exercise 12.5.4 gives rise to an outer measure μ^* defined by $\mu^*(\emptyset) = 0$, $\mu^*(X) = 2$ and $\mu^*(A) = 1$ for any other subset A of X.*

(b): *Now do the construction process again letting $\tau(X) = 3$. What changes?*

Exercise 12.5.6. *Let X be the natural numbers $\mathbb{N}$ and let τ consist of $\emptyset$, $\mathbb{N}$ and all singleton sets. Define $\tau(\emptyset) = 0$ and $\tau(\{x\}) = 1$ for all x in $\mathbb{N}$.*

(a): *Let $\tau(\mathbb{N}) = 2$. Prove the method of Exercise 12.5.4 gives rise to an outer measure μ^*. Determine the family of measurable sets (i.e., the sets that satisfy the Caratheodory Condition).*

(b): *Let $\tau(\mathbb{N}) = \infty$ and answer the same questions as in Part (a).*

(c): *Let $\tau(\mathbb{N}) = 2$ and set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a).*

(d): *Let $\tau(\mathbb{N}) = \infty$ and again set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a). You should see $\mathbb{N}$ is measurable but $\tau(\mathbb{N}) \neq \mu(\mathbb{N})$, where μ denotes the measure constructed in the process of Part (a).*

(e): *Let $\tau(\mathbb{N}) = 1$ and again set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a). What changes?*

Chapter 13

Lebesgue Measure

We will now construct Lebesgue measure on $\Re^k$. We will begin by defining the mapping μ^* on the subsets of $\Re^k$ which will turn out to be an outer measure. The σ - algebra of subsets that satisfy the Caratheodory condition will be called the σ - algebra of Lebesgue measurable subsets. We will denote this σ - algebra by $\mathcal{M}$ as usual. We will usually be able to tell from context what σ - algebra of subsets we are working with in a given study area or problem. The primary references here are again (Bruckner et al. (1) 1997) . and (Taylor (7) 1985) . We like the development of Lebesgue measure in (Taylor (7) 1985) better than that of (Bruckner et al. (1) 1997) and so our coverage reflects that. In all cases, we have added more detail to the proofs of propositions to help you build your analysis skills by looking hard at many interesting and varied proof techniques.

13.1 Lebesgue Outer Measure and Measure

We will be working in $\Re^k$ for any positive integer k. We have to work our way through a fair bit of definitional material; so be patient while we set the stage. We let $x = (x_1, x_2, \ldots, x_k)$ denote a point in the Euclidean space $\Re^k$. An open interval in $\Re^k$ will be denoted by I and it is determined by the cross - product of k intervals of the form (a_i, b_i) where each a_i and b_i is a finite real number. Hence, the interval I has the form

$$I = \Pi_{i=1}^{k} (a_i,\, b_i).$$

The interval (a_i, b_i) is called the i^{th} edge of I and the number $\ell_i = b_i - a_i$ is the length of the i^{th} edge. The **content** of the open interval I is the product of the edge lengths and is denoted by $|I|$; i.e.

$$|I| = \Pi_{i=1}^{k} \Big(b_i - a_i\Big).$$

We need additional terminology. The center of I is the point

$$p = (\frac{a_1 + b_1}{2}, \frac{a_2 + b_2}{2}, \ldots, \frac{a_k + b_k}{2});$$

if the interval J has the same center as the interval I, we say the intervals are concentric.
If I and J are intervals, for convenience of notation, let ℓ_J and ℓ_I denote the vector of edge lengths of J and I, respectively. In general, there is no relationship between ℓ_J and ℓ_I. However, there is a special case of interest. We note that if J is concentric with I and each edge in ℓ_J is a fixed multiple of the corresponding edge length in ℓ_I, we can say $\ell_J = \lambda \ell_I$ for some constant λ. In this case, we write $J = \lambda I$. It then follows that $|J| = \lambda^k \, |I|$.

We are now ready to define outer measure on $\Re^k$. Following Definition 12.3.1, we define a suitable covering family $\mathscr{T}$ and premeasure τ. Then, the mapping μ^* defined in Theorem 12.3.1 will be an outer measure. For ease of exposition, let's define this here.

Definition 13.1.1. Lebesgue Outer Measure

Let $\mathscr{T}$ be the the collection of all open intervals in $\Re^k$ and define the premeasure τ by $\tau(I) = |I|$ for all I in $\mathscr{T}$. For any A in X, define

$$\mu^*(A) \quad = \quad \inf \{ \sum_n |I_n| \,|\, I_n \in \mathscr{T}, \, A \subseteq \cup_n I_n \}$$

We will call a collection (I_n) whose union contains A a Lebesgue Cover of A.

Then, μ^* is an outer measure on $\Re^k$ and as such induces a measure through the usual Caratheodory condition route. It remains to find its properties. The covering family here is not an algebra, so we can not use Theorem 12.3.3 and Theorem 12.3.4 to conclude

(i): $\mathscr{T} \subseteq \mathcal{M}$; i.e. $\mathcal{M}$ is an OMI-F σ - algebra.

(ii): If $A \in \mathscr{T}$, then $|A| = \mu(A)$; i.e. $\mathcal{M}$ is an OMI-FE σ - algebra.

However, we will be able to alter our original proofs to get these results with just a little work.

Comment 13.1.1. *(i): If I is an interval in $\Re^k$, then (I) covers I itself and so by definition $\mu^*(I) \leq |I|$.*

(ii): If $\{x\}$ is a singleton set, choose any open interval I that has x as its center. Then, I is a cover of $\{x\}$ and so $\mu^*(\{x\}) \leq |I|$. We see the the concentric intervals $1/2^n \, I$ also are covers of $\{x\}$ and so $\mu^*(\{x\}) \leq 1/2^n$ for all n. It follows $\mu^*(\{x\}) = 0$.

(iii): From (ii), it clear that $\mu^*(E) = 0$ if E is a finite set.

(iv): If E is countable, label its points by (a_n). Let $\epsilon > 0$ be given. Then by the Infimum Tolerance Lemma, there are intervals I_n having a_n as a center so that $|I_n| < \epsilon/2^n$. Then the intervals (I_n) cover E and by definition,

$$\mu^*(E) \leq \sum_n |I_n| \; \leq \; \sum_n \epsilon/2^n \; = \; \epsilon.$$

Since ϵ is arbitrary, we see $\mu^a st(E) = 0$ if E is countable.

We want to see if $\mu^*(I) = |I|$. This is not clear since our covering family is not an algebra. We now need a technical lemma.

Lemma 13.1.1. Sums Over Finite Lebesgue Covers Of $\overline{I}$ Dominate $|I|$

Let I be any interval of $\Re^k$ and let $(I_1, \ldots, I_N)$ be any finite Lebesgue cover of $\overline{I}$. Then

$$\sum_{n=1}^{N} |I_n| \quad \geq \quad |I|.$$

The proof is based on an algorithm that cycles through the covering sets I_i one by one and picks out certain relevant subintervals. We can motivate this by looking at an interval I in $\Re^2$ whose closure is covered by 3 overlapping intervals I_1, I_2 and I_3. This is shown in Figure 13.1. We do not attempt to indicate the closure of I in this figure nor the fact that the intervals I_1 and so forth are open. We simply draw boxes and you can easily remove or add edges in your mind to open an interval or close it.

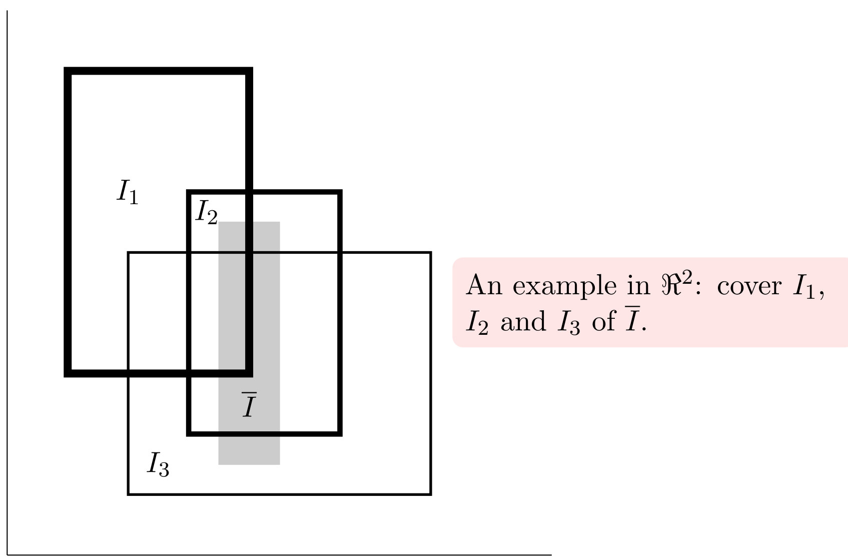


Figure 13.1: *Motivational Lebesgue Cover*

These four intervals all have endpoints on both the x and y axes. If we draw all the possible constant x and constant y lines corresponding to these endpoints, we subdivide the original four intervals into many smaller intervals as shown in Figure 13.2.
In particular, if we looked at interval I_1, it is divided into 16 subintervals (J_1, i), for $1 \leq i \leq 16$ as shown in Figure 13.3.

These rectangles are all disjoint and

$$\overline{I}_1 \quad = \quad \bigcup_{i=1}^{16} \overline{J}_{1,\,i}.$$

although we won't show it in a figure, I_2 and I_3 are also sliced up into smaller intervals; using the same left to right and then downward labeling scheme that we used for I_1, we have

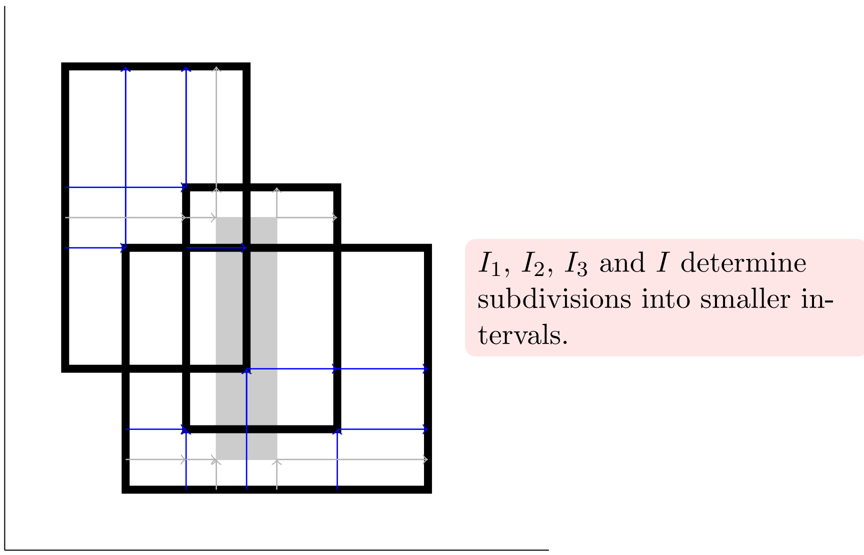


Figure 13.2: *Subdivided Lebesgue Cover*

- I_2 is divided by 4 horizontal and 4 vertical lines into 16 disjoint subintervals, $J_{2,1}$ to $J_{2,16}$. Further,

$$\overline{I}_2 \quad = \quad \bigcup_{i=1}^{16} \overline{J}_{2,\,i}.$$

- I_3 is divided by 4 horizontal and 6 vertical lines into 24 disjoint subintervals, $J_{3,1}$ to $J_{3,24}$. We thus know

$$\overline{I}_3 \quad = \quad \bigcup_{i=1}^{24} \overline{J}_{3,\,i}.$$

Finally, I is also subdivided into subintervals: it is divided by 4 horizontal and 2 vertical lines into 8 disjoint subintervals, J_1 to J_8 and

$$\overline{I} \quad = \quad \bigcup_{i=1}^{8} \overline{J}_i.$$

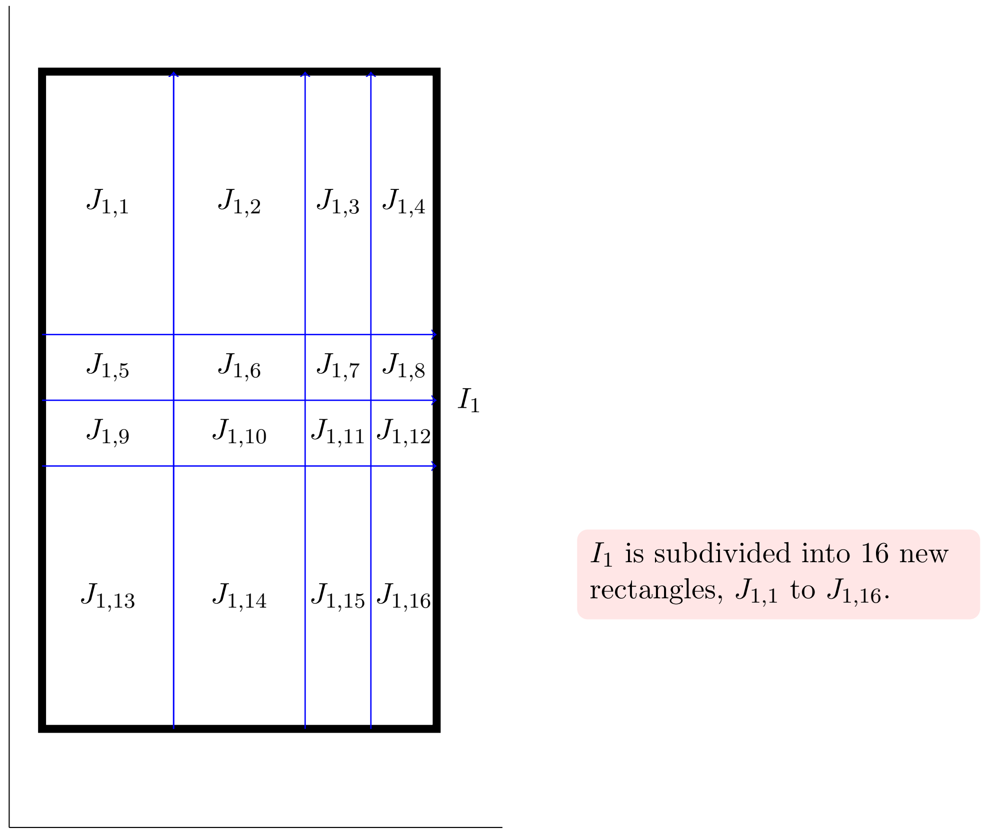


Figure 13.3: *Subdivided* I_1

We also know

$$\begin{aligned}
|I| &= \sum_{i=1}^{8} |J_i|, \\
|I_1| &= \sum_{i=1}^{16} |J_{1,i}|, \\
|I_2| &= \sum_{i=1}^{16} |J_{2,i}|, \\
|I_3| &= \sum_{i=1}^{24} |J_{3,i}|.
\end{aligned}$$

Now look at Figure 13.2 and you see immediately that the intervals J_{kj} and J_{pq} are either the same or are disjoint. For example, the subintervals match when interval I_2 and I_3 overlap. We can conclude each J_i is disjoint from a J_{kj} or it equals J_{kj} for some choice of k and j. Here is the algorithm we want to use:

Step 1: We know $I \subseteq I_1 \cup I_2 \cup I_3$ and $J_1 = J_{n_1,q_1}$ where n_1 is the smallest index from 1, 2 or 3 which works. For this fixed n_1, consider the collection

$$S_{n_1} \quad = \quad \{J_{n_1,1},\ \ldots\ , J_{n_1,p(n_1)}\}$$

where we are using the symbol $p(n_1)$ to denote the number of subintervals for I_{n_1}. Thus, $p(1) = p(2) = 16$ and $p(3) = 24$ in our example. In our example, we find $n_1 = 1$ and

$$\begin{aligned} J_1 &= J_{1,12} \\ S_1 &= \{J_{1,1},\ \ldots\ , J_{1,16}.\} \end{aligned}$$

Look at Figure 13.4 to see what we have done so far.

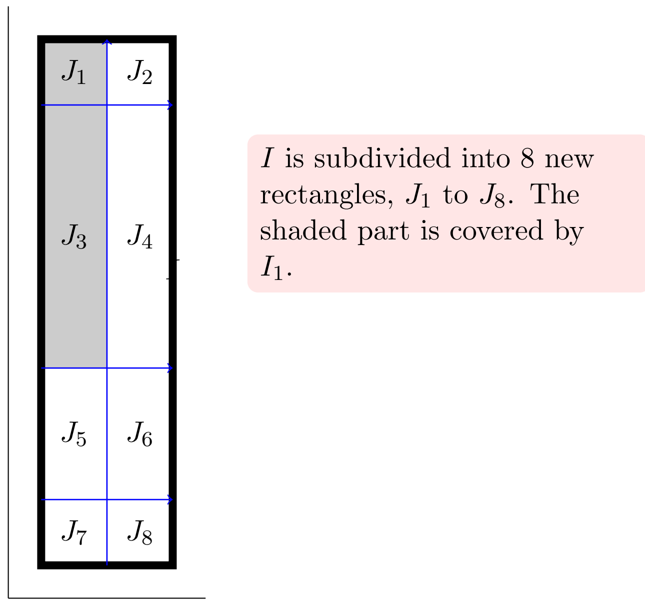


Figure 13.4: *The Part Of I Covered by I_1*

By referring to Figure 13.2, you can see $J_1 = J_{1,12}$ and $J_3 = J_{1,16}$. Now, let

$$T_{n_1} \quad \equiv \quad T_1 \;=\; \{i \,|\, \exists k \,\ni\, J_i \;=\; J_{n_1,k}\}.$$

Here $T_1 = \{1,\ 3\}$. Also, let

$$U_{n_1} \quad \equiv \quad U_1 \;=\; \{k \,|\, \exists i \,\ni\, J_{n_1,k} \;=\; J_i\}.$$

We see $U_1 = \{12,\ 16\}$.

Step 2: Now look at the indices

$$\begin{aligned} V_{n_1} \quad &\equiv \quad V_1 \;=\; \{1, 2, 3, \ldots, 8\} \setminus T_1 \\ &= \quad \{2, 4, 5, 6, 7, 8\}. \end{aligned}$$

The smallest index in this set is 2. Next, find the smallest index n_2 so that

$$J_2 \quad = \quad J_{n_2,k}$$

for some index k. From Figure 13.2, we see both I_2 and I_3 intersect $I \setminus I_1$. The smallest index n_2 is thus $n_2 = 2$. The index k that works is 7 and so $J_2 = J_{2,7}$. In figure 13.5, we have now shaded the part of I not in I_1 that lies in I_2.

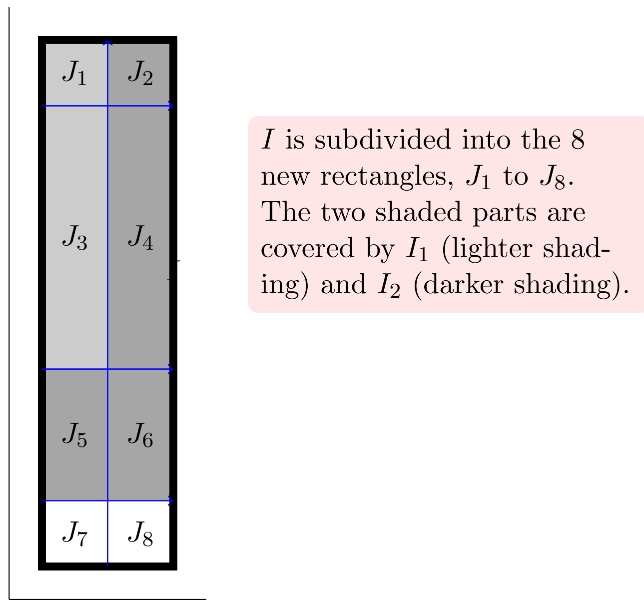


Figure 13.5: *The Part Of I Covered by I_1 and I_2*

We can see that $J_2 = J_{2,7}$, $J_4 = J_{2,11}$, $J_5 = J_{2,14}$ and $J_6 = J_{2,15}$. Let

$$T_{n_2} \quad \equiv \quad T_2 \; = \; \{i \in V_1 \,|\, \exists k \ni \; J_i \; = \; J_{n_2,k}\}.$$

Here $T_2 = \{2,\ 4,\ 5,\ 6\}$. Also, let

$$U_{n_2} \quad \equiv \quad U_2 \; = \; \{k \,|\, \exists i \ni \; J_{n_2,k} \; = \; J_i\}.$$

We see $U_1 = \{7,\ 11,\ 14,\ 15\}$.

Step 3: Now look at the indices

$$\begin{aligned} V_{n_2} \quad &\equiv \quad V_2 \; = \; \{1,2,3,\ldots,8\} \setminus (T_1 \cup T_2) \\ &= \quad \{7,8\}. \end{aligned}$$

The smallest index in this set is 7. Next, find the smallest index n_3 so that

$$J_7 \quad = \quad J_{n_3,k}$$

for some index k. From Figure 13.2, we see both I_2 and I_3 intersect $I \setminus (I_1 \cup I_2)$. The smallest index n_3 must be 3 and so $n_3 = 3$. The index k that works now is 15 and we have $J_7 = J_{3,15}$. In figure 13.6, we have now shaded the part of I not in $I_1 \cup I_2$ that lies in I_3.

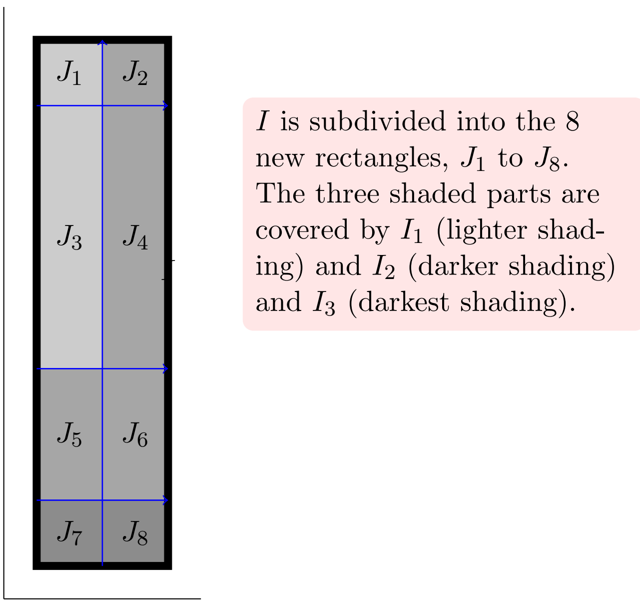


Figure 13.6: *The Part Of I Covered by I_1, I_2 and I_3*

In fact, we have $J_7 = J_{3,15}$ and $J_8 = J_{3,16}$. Thus, we set

$$\begin{aligned} T_{n_3} \;&\equiv\; T_3 \;=\; \{i \in V_2 \,|\, \exists k \ni J_i \;=\; J_{n_3,k}\} \\ &=\; \{7,\, 8\}. \end{aligned}$$

Also, we let

$$U_{n_3} \;\equiv\; U_3 \;=\; \{k \,|\, \exists i \ni J_{n_3,k} \;=\; J_i\}.$$

We see $U_1 = \{15,\, 16\}$.

We have now expressed each J_i as some $J_{n_1,k}$ through $J_{n_3,k}$. We are now ready to finish our argument.

Step 4: We have

$$\begin{aligned} \{1,\, \ldots,\, 8\} \;&=\; T_{n_1} \cup T_{n_2} \cup T_{n_3} \\ &=\; T_1 \cup T_2 \cup T_3. \end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{k \in U_{n_3} = U_3} |J_{n_3,k}| &\leq \sum_{k=1}^{p(n_3)} |J_{n_3,k}| = \sum_{k=1}^{24} |J_{3,k}| \leq |I_3|, \\
\sum_{k \in U_{n_2} = U_2} |J_{n_2,k}| &\leq \sum_{k=1}^{p(n_2)} |J_{n_2,k}| = \sum_{k=1}^{16} |J_{2,k}| \leq |I_2|, \\
\sum_{k \in U_{n_1} = U_1} |J_{n_1,k}| &\leq \sum_{k=1}^{p(n_1)} |J_{n_1,k}| = \sum_{k=1}^{16} |J_{1,k}| \leq |I_{n_1}|.
\end{aligned}$$

Thus,

$$\begin{aligned}
|I| &= \sum_{i=1}^{8} |J_i| = \sum_{p=1}^{3} \sum_{k \in U(n_p)} |J_{n_p,k}| \\
&\leq \sum_{p=1}^{3} |I_{n_p}|.
\end{aligned}$$

This proves that

$$|I| \leq \sum_{i=1}^{3} |I_i|.$$

This is our desired proposition for a particular example set in $\Re^2$ using three intervals. We are now ready to adapt this algorithm to prove the general result.

Proof. *We are given intervals I_1 to I_N in $\Re^k$ whose union covers $\overline{I}$. Each interval I_i is the product*

$$(\alpha_{i1}, \beta_{i1}) \times \cdots \times (\alpha_{ik}, \beta_{ik}),$$

and I is the product

$$(\alpha_1, \beta_1) \times \cdots \times (\alpha_k, \beta_k).$$

On the x_j axis, the N intervals and the interval I determine a collection of points

$$\begin{aligned}
\{(\alpha_{1j}, \beta_{1j}),\ x_j &\ \textit{edge from interval}\ I_1; \\
(\alpha_{2j}, \beta_{2j}),\ x_j &\ \textit{edge from interval}\ I_2; \\
&\ \vdots \\
(\alpha_{Nj}, \beta_{Nj}),\ x_j &\ \textit{edge from interval}\ I_N; \\
(\alpha_j, \beta_j),\ x_j &\ \textit{edge from interval}\ I.
\end{aligned}$$

We do not care if these points are ordered. These x_j axis points, for $1 \leq j \leq k$, "slice" the intervals I_1 through I_N and I into smaller intervals just as we did in the example for $\Re^2$ shown in Figure 13.2.

We have

$$
\begin{array}{ccl}
I & \longrightarrow & J_1, \ldots, J_p \\
I_1 & \longrightarrow & J_{11}, \ldots, J_{1,p(1)} \\
\vdots & & \\
I_N & \longrightarrow & J_{N1}, \ldots, J_{N,p(1)}.
\end{array}
$$

Step 1: *Look at* J_1*. There is a smallest index* n_1 *so that* $J - 1 = J_{n_1,\ell}$ *for some* ℓ*. Let*

$$
\begin{aligned}
T_{n_1} &= \{\, i\{1,\ldots,p\} \,|\, \exists\, \ell \ni J_i = J_{n_1,\ell}\}, \\
U_{n_1} &= \{\, \ell \,|\, \exists\, i \ni J_i = J_{n_1,\ell}\}.
\end{aligned}
$$

This uses up T_{n_1} *of the indices* $\{1,\ldots,p\}$*. You can see this process in Figure 13.4.*

Step 2: *Let*

$$
V_1 = \{1, \ldots, p\} \setminus T_{n_1}\}
$$

and let q *be the smallest index from the set* V_1*. For this* q*, find the smallest index* $n_2 \neq n_1$ *so that* $J_q = J_{n_2,\ell}$ *for some* ℓ*. This is the process we are showing in Figure 13.5. We define*

$$
\begin{aligned}
T_{n_2} &= \{\, i \in V_1 \,|\, \exists\, \ell \ni J_i = J_{n_2,\ell}\}, \\
U_{n_2} &= \{\, \ell \,|\, \exists\, i \in V_1 \ni J_i = J_{n_2,\ell}\}.
\end{aligned}
$$

This uses up more of the smaller subintervals I_1 *to* I_p*.*

Additional Steps : *Let*

$$
V_2 = \{1, \ldots, p\} \setminus (T_{n_1} \cup T_{n_2})\}.
$$

We see V_2 *is a smaller subset of the original* $\{1, \ldots, p\}$ *than* V_1*. We continue this construction process until we have used up all the indices in* $\{1, \ldots, p\}$*. This takes say* Q *steps and we know* $Q \leq p$*.*

Final Step: *After the process terminates, we have*

$$
\begin{aligned}
|I| &= \sum_{i=1}^{p} |J_i| \\
&= \sum_{p=1}^{Q} \sum_{\ell \in U(n_p)} |J_{n_p,\ell}| \\
&\leq \sum_{p=1}^{Q} |I_{n_p}| \leq \sum_{i=1}^{N} |I_i|.
\end{aligned}
$$

this completes the proof. ■

We can now finally prove that $\mu^*(I) = |I|$. Note that we have to work this hard because our original covering family was not an algebra! The final arguments are presented in the next two lemmatta.

Lemma 13.1.2. $\mu^*(\overline{I} \ = \ |I|$

Let I be an open interval in $\Re^k$. Then $\mu^*(\overline{I}) \ = \ |I|$.

Proof. *Let (I_n) be any Lebesgue cover of $\overline{I}$. Since $\overline{I}$ is compact, this cover has a finite subcover, $I_{n_1}, \ldots, I_{n_N}$. Applying Lemma 13.1.1, we see*

$$|I| \ \leq \ \sum_{i=1}^{N} |I_{n_i}| \leq \sum_i |I_i|.$$

Since (I_n) is an arbitrary cover of $\overline{I}$, we then have $|I|$ is a lower bound for the set

$$\{\sum_n |I_n| \,|\, (I_n) \ \text{is a cover of} \ \overline{I}\}.$$

It follows that

$$|I| \ \leq \ \mu^*(\overline{I}).$$

To prove the reverse inequality holds, let U be an open interval concentric with I so that $\overline{I} \subseteq U$. Then U is a cover of $\overline{I}$ and so $\mu^*(\overline{I}) \ \leq \ |U|$. Hence, for any concentric interval, λI, $1 < \lambda < 2$, we have $\mu^*(\overline{I}) \ \leq \ \lambda^k \, |I|$. Since this holds for all $\lambda > 1$, we can let $\lambda \rightarrow 1$ to obtain $\mu^*(\overline{I}) \ \leq |I|$. ■

Lemma 13.1.3. $\mu^*(I) \ = \ |I|$

If I is an open interval of $\Re^k$, then $\mu^*(I) \ = \ |I|$.

Proof. *We know I is a cover of itself, so it is immediate that $\mu^*(I) \leq |I|$. To prove the reverse inequality, let λI be concentric with I for any $0 < \lambda < 1$. Then, $\overline{\lambda I} \subseteq I$ and since μ^* is an outer measure, it is monotonic and so*

$$\mu^*(\overline{\lambda I}) \ \leq \ \mu^*(I).$$

But $\mu^*(\overline{\lambda I}) = \lambda^k \, |I|$. We thus have $\lambda^k \, |I| \leq \mu^*(I)$ for all $\lambda \in (0,1)$. Letting $\lambda \rightarrow 1$, we obtain the desired inequality. ■

13.2 Lebesgue Outer Measure Is A Metric Outer Measure

We have now shown that if $I \in \mathscr{T}$, then $|I| = \mu^*(I)$. However, we still do not know that the intervals I from $\mathscr{T}$ are μ^* measurable. We will do this by showing that Lebesgue outer measure is a metric outer measure. Then, it will follow from Theorem 12.2.1 that the open sets in $\Re^k$ are μ^* measurable, i.e. are in $\mathcal{M}$. Of course, this implies $\mathscr{T} \subseteq \mathcal{M}$ as well. Then, since an interval I is measurable, we have $|I| = \mu(I)$. Let's prove μ^* is a metric outer measure. We begin with a technical definition.

Definition 13.2.1. The M_δ Form of μ^*

For any set E is $\Re^k$ and any $\delta > 0$, let

$$M_\delta(E) \quad = \quad \inf \{ \sum_n |I_n| \, | \, (I_n) \ \text{covers } E, \, I_n \ \text{is an interval } \in R^k, \ \text{each edge of } I_n \ \text{is less than } \delta \}.$$

Next, we need a technical lemma concerning finite Lebesgue covers.

Lemma 13.2.1. Approximate Finite Lebesgue Covers Of $\overline{I}$.

Let I be a open interval and let $\overline{I}$ denote its closure. Let ϵ and δ be given positive numbers. Then there exists a finite Lebesgue Covering of $\overline{I}$, $I_1, \ldots, I_N$ so that each edge of I_i has length less than δ and

$$|I_1| + \cdots + |I_N| \quad < \quad |I| + \epsilon$$

Proof. *Let*

$$I \quad = \quad \Pi_{i=1}^k (a_i, b_i)$$

and divide each component interval (a_i, b_i) into n_i uniform pieces so that $(b_i - a_i)/2 < \delta/2$. This determines n_i open intervals of the form (a_{ij}, b_{ij}) for $1 \leq j \leq n_i$ with $b_{ij} - a_{ij} < \delta/2$.

Let $N = n_1 \, n_2 \, \cdot \, n_k$ and let $J = (j_1, \ldots, j_k)$ denote the k - tuple of indices chosen so that $1 \leq j_i \leq n_i$. There are N of these indices. Let $\boldsymbol{j}$ indicate any such k - tuple. Then $\boldsymbol{j}$ determines an interval $I_{\boldsymbol{j}}$ where

$$I_{\boldsymbol{j}} \quad = \quad \Pi_{i=1}^k (a_{ij}, b_{ij}), \quad \text{with } \ (b_{ij} - a_{ij}) < \delta/2.$$

Hence, $|I_{\boldsymbol{j}}| < (\delta/2)^k$. It is also clear that

$$\sum |I_{\boldsymbol{j}}| \quad = \quad |I|.$$

Now choose concentric open intervals $\lambda I_{\boldsymbol{j}}$ for any λ with $1 < \lambda < 2$. Then since $\lambda > 1$, $(\lambda I_{\boldsymbol{j}}$ over all k - tuples $\boldsymbol{j}$ is a Lebesgue cover of $\overline{I}$, we have

$$|\lambda I_{\boldsymbol{j}}| \quad = \quad \lambda^k \, |I_{\boldsymbol{j}}|$$

and so

$$\begin{aligned} \sum |\lambda I_{\boldsymbol{j}}| \quad &= \quad \lambda^k \sum |I_{\boldsymbol{j}}| \\ &= \quad \lambda^k \, |I|. \end{aligned}$$

Since $\lambda^k \to 1$, for our given $\epsilon > 0$, there is a $\eta > 0$ so that if $1 < \lambda < 1 + \eta$, we have

$$\lambda^k \, - \, 1 \quad < \quad \frac{\epsilon}{|I| \, + \, 1}.$$

In particular, if we pick $\lambda = (1+\eta)/2$, then

$$|\lambda I_{\boldsymbol{j}}| \quad < \quad \left(1 \,+\, \frac{\epsilon}{|I| \,+\, 1}\right) |I| \,<\, |I| \,+\, \epsilon.$$

Since ϵ is arbitrary, we see

$$|\lambda I_{\boldsymbol{j}}| \quad < \quad \left(1 \,+\, \frac{\epsilon}{|I| \,+\, 1}\right) |I| \,<\, |I| \,+\, \epsilon.$$

Thus, the finite collection $((1+\eta)/2\, I_{\boldsymbol{j}})$ is the one we seek as each edge has length $((1+\eta)/2\, \delta/2$ which is less than δ. ■

Lemma 13.2.2. $M_\delta \;=\; \mu^*$

For any subset E of $\Re^k$, we have $M_\delta(E) \;=\; \mu^*(E)$.

Proof. *Let's pick a given $\delta > 0$. The way M_δ is defined then tells us immediately that $\mu^*(E) \leq M_\delta(E)$ for any $\delta > 0$ and subset E. It remains to prove the reverse inequality. If $\mu^*(E)$ was infinite, we would have $\mu^*(E) \geq M_\delta(E)$; hence, it is enough to handle the case where $\mu^*(E)$ is finite. By the Infinitum Tolerance Lemma for a given $\epsilon > 0$, there is a Lebesgue cover (I_n) of E so that*

$$\sum_n |I_n| \quad < \quad \mu^*(E) \,+\, \frac{\epsilon}{2}.$$

By Lemma 13.2.1, there is a finite Lebesgue cover of each $(\overline{I_n})$ which we will denote by (J_{nj}), $1 \leq j \leq p(n)$ so that each interval J_{nj} has edge length less than δ and satisfies

$$\sum_{j=1}^{p(n)} |J_{nj}| \quad < \quad |I_n| \,+\, \frac{\epsilon}{2^{n+1}}.$$

The combined family of intervals (J_{nj} for all n and $1 \leq j \leq p(n)$ is clearly a Lebesgue cover of E also. Thus, by definition of μ^*, we have

$$\begin{aligned} \sum_{n=1}^{\infty} \sum_{j=1}^{p(n)} |J_{nj}| \quad &< \quad \sum_{n=1}^{\infty} |I_n| \,+\, \sum_{n=1}^{\infty} \frac{\epsilon}{2^{n+1}} \\ &< \quad \mu^*(E) \,+\, \epsilon. \end{aligned}$$

Now each edge length of the interval I_{nj} is less than δ and so

$$M_\delta \quad \leq \quad \sum_{n=1}^{\infty} \sum_{j=1}^{p(n)} |J_{nj}|$$

by definition. We see we have established

$$M_\delta \quad \leq \quad \mu^*(E) \,+\, \epsilon$$

for an arbitrary ϵ; hence, $M_\delta \leq \mu^*(E)$. ■

We now have enough "ammunition" to prove Lebesgue outer measure is a metric outer measure; i.e. LOM is a MOM!

Theorem 13.2.3. Lebesgue Outer Measure Is a Metric Outer Measure
The Lebesgue Outer Measure, μ^* is a metric outer measure; i.e., if A and B are two sets in $\Re^k$ with $D(A,B) > 0$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Proof. *We always know that $\mu^*(A \cup B) \leq \mu^*(A) + \mu^*(B)$ for any A and B. Hence, for two sets A and B with $D(A,B) = \delta > 0$, it is enough to show $\mu^*(A) + \mu^*(B) \leq \mu^*(A \cup B)$. Let $\epsilon > 0$ be chosen. Since $M_\delta = \mu^*$, there is a cover of $A \cup B$ so that the edge length of each I_n is less than δ/k and*

$$M_\delta(A \cup B) \;=\; \mu^*(A \cup B) \leq \sum_n |I_n| < \mu^*(A \cup B) + \epsilon$$

by an application of the Infimum Tolerance Lemma.

If x and y in $A \cup B$ are both in a given I_n, then

$$d(x,y) \;=\; \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} < \sqrt{\sum_{i=1}^{k} (\frac{\delta}{k})^2} = \sqrt{k^2 \frac{\delta^2}{k^2}} = \delta.$$

However, $D(A,B) = \delta$ by assumption. Thus, a given I_n can not contain points of both A and B. We can therefore separate the family (I_n) into two collections indexed by U and V, respectively. If $n \in U$, then $I_n \cap A$ is non empty and if $n \in V$, $I_n \cap B$ is non empty. We see $\{I_n\}_{n \in U}$ is a cover for A and $\{I_n\}_{n \in V}$ is a cover for B. Thus, $\mu^*(A) \leq \sum_{n \in U} |I_n|$ and $\mu^*(B) \leq \sum_{n \in V} |I_n|$. It then follows that

$$\begin{aligned} \mu^*(A \cup B) + \epsilon &\geq \sum_n |I_n| = \sum_{n \in U} |I_n| + \sum_{n \in V} |I_n| \\ &\geq \mu^*(A) + \mu^*(B). \end{aligned}$$

Since ϵ is arbitrary, we have shown $\mu^*(A) + \mu^*(B) \leq \mu^*(A \cup B)$. This completes the proof that Lebesgue outer measure is a metric outer measure. ■

This theorem is the final piece we need to fully establish the conditions

(i): $\mathscr{T} \subseteq \mathcal{M}$; i.e. $\mathcal{M}$ is an OMI-F σ - algebra.

(ii): If $I \in \mathscr{T}$, then $|I| = \mu(I)$; i.e. $\mathcal{M}$ is an OMI-FE σ - algebra.

Comment 13.2.1. *We see immediately that since Lebesgue outer measure is a metric outer measure, the σ - algebra of μ^* measurable subsets contains all the open sets of $\Re^k$. In particular, any open interval I is measurable. As mentioned previously, we thus know the Borel σ - algebra of subsets is contained in $\mathcal{M}$.*

By Theorem 12.1.4, we know Lebesgue measure μ is complete.

We can also prove Lebesgue measure μ is regular.

Theorem 13.2.4. Lebesgue Measure Is Regular

For any set E in $\Re^k$,

$$\begin{aligned}\mu^*(E) &= \inf \{\mu(U) \,|\, U,\ E \subseteq U,\ U \text{ is open } \}\\ \mu^*(E) &= \inf \{\mu(F) \,|\, E,\ E \subseteq F,\ F \text{ is Lebesgue measurable } \}.\end{aligned}$$

Hence, μ is regular.

Proof. *Since U is open, U is Lebesgue measurable and so $\mu^*(U) = \mu(U)$. It follows immediately that $\mu^*(E) \leq \mu(U)$ for such U. Hence,*

$$\mu^*(E) \leq \inf \{\mu(U) \,|\, U,\ E \subseteq U,\ U \text{ is open } \}.$$

On the other hand, if $\epsilon > 0$ is given, the Infimum Tolerance Lemma tells us there is a Lebesgue cover of E, (I_n), so that

$$\mu^*(E) \leq \sum_n |I_n| < \mu^*(E) + \epsilon.$$

However, this open cover generates an open set $G = \cup_n I_n$ containing E with $\mu(G) \leq \sum_n |I_n|$ because $\mu(I_n) = |I_n|$. We conclude, using the definition of μ^* that

$$\mu(G) \leq \sum_n |I_n| < \mu^*(E) + \epsilon.$$

Hence, we must have

$$\inf \{\mu(U) \,|\, U,\ E \subseteq U,\ U \text{ is open } \} \leq \mu^*(E) + \epsilon.$$

Since ϵ is arbitrary, the result follows.

Since each open U is measurable, we then know

$$\begin{aligned}\mu^*(E) &= \inf \{\mu(U) \,|\, U,\ E \subseteq U,\ U \text{ is open } \}\\ &\geq \inf \{\mu(F) \,|\, E,\ E \subseteq F,\ F \in \mathcal{M}\}\end{aligned}$$

by the first argument. To obtain the reverse inequality, note that since $\mu^*(F) = \mu(F)$ for all measurable F, monotonicity of μ^* says $\mu^*(E) \leq \mu^*(F)$ for all measurable F. We conclude

$$\mu^*(E) \leq \inf \{\mu(F) \,|\, E,\ E \subseteq F,\ F \in \mathcal{M}\}.$$

Now recall the definition of a regular measure from Definition 12.3.3. Using the Infimum Tolerance Lemma again, there is are measurable sets (F_n) so that $E \subseteq F_n$ for all n and

$$\mu^*(E) \leq \mu(F_n) < \mu^*(E) + \frac{1}{n}.$$

Then, $\cap_n F_n$ is also measurable and so by our equivalent form of μ^*, we have $\mu^*(E) \leq \mu(\cap_n F_n)$. However, $\cap_n F_n \subseteq F_n$ always and hence,

$$\mu^*(E) \quad \leq \quad \mu(\cap_n F_n) \;\leq\; \mu(F_n) \;<\; \mu^*(E) \;+\; \frac{1}{n}.$$

We conclude for all n,

$$\mu^*(E) \quad \leq \quad \mu(\cap_n F_n)\mu^*(E) \;+\; \frac{1}{n}.$$

Letting n go to infinity, we find $\mu^*(E) = \mu(\cap_n F_n)$ which shows μ is regular. ■

13.3 Lebesgue - Stieljes Outer Measure and Measure

We can also be more general. Let g be any non -decreasing function on $\Re$ which is continuous from the right. This means for all x, $\lim_{h \rightarrow 0^+} g(x+h)$ exists. Moreover, the unbounded limits are well - defined $\lim_{x \rightarrow -\infty} g(x)$ and $\lim_{x \rightarrow \infty} g(x)$. These last two limits could be $-\infty$ and ∞ respectively. Then, define the mapping τ_g on $\mathcal{U}$ by

$$\begin{aligned}
\tau_g(\emptyset) &= 0, \\
\tau_g\Bigl((a,b]\Bigr) &= g(b) \;-\; g(a), \\
\tau_g\Bigl((-\infty,b]\Bigr) &= g(b) \;-\; \lim_{x \rightarrow -\infty} g(x), \\
\tau_g\Bigl((a,\infty)\Bigr) &= \lim_{x \rightarrow \infty} g(x) \;-\; g(a), \\
\tau_g\Bigl((-\infty,\infty)\Bigr) &= \lim_{x \rightarrow \infty} g(x) \;-\; \lim_{x \rightarrow -\infty} g(x).
\end{aligned}$$

This defines τ_g on the collection of sets $\mathcal{U}$ consisting of the empty set, intervals of the form $(a, b]$ for finite numbers a and b and unbounded intervals of the form $(-\infty, b]$ and (a, ∞). Let $\mathcal{A}$ be the algebra generated by finite unions of sets from $\mathcal{U}$. Note $\mathcal{A}$ contains $\Re$.

Let's extend the mapping τ_g to be additive on $\mathcal{A}$. If $E_1, E_2, \ldots, E_n$ is a finite collection of disjoint sets in $\mathcal{A}$, we extend the definition of τ_g to this finite disjoint unions as follows:

$$\tau_g(\cup_{i-1}^n E_i) \quad = \quad \sum_{i=1}^n \tau_g(E_i). \tag{13.1}$$

Lemma 13.3.1. Extending τ_g To Additive Is Well - Defined
The extension of τ_g from $\mathcal{U}$ to the algebra $\mathcal{A}$ is well - defined; hence, τ_g is additive on $\mathcal{A}$.

Proof. *For $(a, b] \in \mathcal{A}$, write*

$$(a,b] \quad = \quad \cup_{i=1}^n (a_i, b_i],$$

for any positive integer n with $a_1 = a$, $b_n = b$ and the in between points satisfy $a_{i+1} = b_i$ for all i. Of course, there are many such decompositions of $(a, b]$ we could choose. Also, these are the only decompositions we can have. If we use the unbounded sets, we can not recapture $(a, b]$ using a finite number of unions! Then, using Equation 13.1, we have

$$\begin{aligned}\tau_g((a,b]) &= \sum_{i=1}^{n} \tau_g((a_i, b_i]) \\ &= \sum_{i=1}^{n} g(b_i) \; - \; g(a_i).\end{aligned}$$

But since $a_{i+1} = b_i$, this sum collapses to

$$\tau_g((a,b]) \quad = \quad g(b) \; - \; g(a).$$

This was the original definition of τ_g on the element $(a, b]$ in $\mathcal{U}$. We conclude the value of τ_g on elements of the form $(a, b]$ is independent of the choice of decomposition of it into a finite union of sets from $\mathcal{U}$.

For an unbounded interval of the form (a, ∞), any finite disjoint decomposition can have only one interval of the form (b, ∞) giving $(a, \infty) = (a, b] \cup (b, \infty)$, with the piece $(a, b]$ written as any finite disjoint union $(a, b] = \cup_{i=1}^{n} (a_i, b_i]$ as before. The same arguments as used above then show τ_g is well - defined on this type of element of $\mathcal{U}$ also. We handle the sets $(-\infty, b]$ is a similar fashion.

Next, if we look at any arbitrary A in $\mathcal{A}$, then A can be written as a finite union of members $A_1, \ldots, A_p$ of $\mathcal{U}$. Each of these elements A_i can then be written using a finite disjoint decomposition into intervals $(a_{ij}, b_{ij}]$, $1 \leq j \leq p(i)$ as we have done above. Thus,

$$A \quad = \quad \cup_{i=1}^{m} \cup_{j=1}^{p(i)} \; (a_{ij}, b_{ij}]$$

where it is possible $a_{11} = -\infty$ and $b_{m\,p(m)} = \infty$. We then combine these intervals and relabel as necessary to write A as a finite disjoint union

$$A \quad = \quad \cup_{i=1}^{N} (a_i, b_i]$$

with $b_i \leq a_{i+1}$ and again it is possible that $a_1 = -\infty$ and $b_N = \infty$. We therefore know that

$$\tau_g(A) \quad = \quad \cup_{i=1}^{N} \tau_g((a_i, b_i]).$$

Now assume A has been decomposed into another finite disjoint union, $A = \cup_{j=1}^{M} B_j$, each $B_j \in \mathcal{A}$. Let

$$C_j \quad = \quad \{i \mid \; \subseteq (a_i, b_i] \; \subseteq \; B_j\}.$$

Note a given interval $(a_i, b_i]$ can not be in two different sets B_j and B_k because they are assumed disjoint. Hence, we have

$$B_j \quad = \quad \cup_{i \in C_j} (a_i, b_i]$$

and

$$\tau_g(B_j) \quad = \quad \sum_{i \in C_j} \tau_g((a_i, b_i]).$$

Thus,

$$\begin{aligned}\sum_{j=1}^{M} \tau_g(B_j) &= \sum_{j=1}^{M} \sum_{i \in C_j} \tau_g((a_i, b_i]) \\ &= \sum_{i=1}^{N} \tau_g((a_i, b_i]).\end{aligned}$$

This shows that our extension for τ_g is independent of the choice of finite decomposition and so the extension of τ_g is a well - defined additive map on $\mathcal{A}$ ■

We can now apply Theorem 12.3.3 to conclude that since the covering family $\mathscr{A}$ is an algebra and τ_g is additive on $\mathscr{A}$, the σ - algebra, $\mathcal{M}_g$, generated by τ_g contains $\mathscr{A}$ and the induced measure, μ_g, is regular. Next, we want to know that $\mu_g(A) = \tau_g(A)$ for all A in $\mathcal{A}$. To do this, we will prove the extension τ_g is actually a pseudo-measure. Thus, we will be able to invoke Theorem 12.3.4 to get the desired result.

Lemma 13.3.2. Lebesgue - Stieljes Premeasure Is a Pseudo-Measure

The mapping τ_g is a pseudo-measure on $\mathcal{A}$.

Proof. *We need to show that if (T_n) is a sequence of disjoint sets from $\mathcal{A}$ whose union $\cup_n T_n$ is also in $\mathcal{A}$, then*

$$\tau_g(\cup_n T_n) \quad = \quad \sum_n \tau_g(T_n).$$

First, notice that if there was an index n_0 so that $\tau_g(T_{n_0}) = \infty$, then letting $B = \cup_n T_n \setminus T_{n_0}$, we can write $\cup_n T_n$ as the finite disjoint union $B \cup T_{n_0}$ and hence

$$\tau_g(\cup_n T_n) \quad = \quad \tau_g(B) + \tau_g(T_{n_0}) = \infty.$$

Since the right hand side sums to ∞ in this case also, we see there is equality for the two expressions. Therefore, we can restrict our attention to the case where all the individual T_n sets have finite $\tau_g(T_n)$ values. This means no elements of the form $(-\infty, b]$ or (a, ∞) can be part of any decomposition of the sets T_n. Hence, we can assume each T_n can be written as a finite union of intervals of the form $(a, b]$. It follows then that it suffices to prove the result for a single interval of the form $(a, b]$.

Since τ_g is additive on finite unions, if $C \subseteq D$, we have

$$\tau_g(D) \quad = \quad \tau_g(C) + \tau_g(D \setminus C) \geq \tau_g(C).$$

Now assume we can write the interval (a, b) as follows:

$$(a, b] \quad = \quad \cup_{n=1}^{\infty} (a_i, b_i]$$

with the sets $(a_i, b_i]$ disjoint. For any n, we have

$$(a, b] \quad = \quad \cup_{k=1}^{n} (a_k, b_k] \, \cup \, \cup_{k=n+1}^{\infty} (a_k, b_k].$$

Therefore

$$\tau_g\biggl((a, b] \biggr) \quad = \quad \tau_g\biggl(\cup_{k=1}^{n} (a_k, b_k] \biggr) \, + \, \tau_g\biggl(\cup \cup_{k=n+1}^{\infty} (a_k, b_k] \biggr).$$

The finite additivity on disjoint intervals then gives us

$$\begin{aligned} \tau_g\biggl(\cup_{k=1}^{n} (a_k, b_k] \biggr) \quad &= \quad \sum_{k=1}^{n} \tau_g\biggl((a_k, b_k] \biggr) \\ &= \quad g(b_1) \, - \, g(a_1) \, + \, g(b_2) \, - \, g(a_2) \, + \, \ldots \, + \, g(b_n) \, - \, g(a_a). \end{aligned}$$

We know g is nondecreasing, thus $g(b_1) - g(a_2) \leq 0$, $g(b_2) - g(a_3) \leq 0$, and so forth until we reach $g(b_{n-1}) - g(a_n) \leq 0$. Dropping these terms, we find

$$\tau_g\biggl(\cup_{k=1}^{n} (a_k, b_k] \biggr) \quad \leq \quad g(b_n) \, - \, g(a_1) \, \leq g(b) \, - \, g(a).$$

Thus, these partial sums are bounded above and so the series of non negative terms $\sum_n \tau_g((a_k, b_k])$ converges. This tells us that

$$\tau_g\biggl(\cup_{k=1}^{\infty} (a_k, b_k] \biggr) \quad \leq \quad \tau_g\biggl((a, b] \biggr).$$

To obtain the reverse inequality, let $\epsilon > 0$ be given. Then, since the series above converges, there must be a positive integer N so that if $n \geq N$,

$$\sum_{k=n+1}^{\infty} \tau_g\biggl((a_k, b_k] \biggr). \quad < \quad \epsilon$$

We conclude that

$$\begin{aligned} \tau_g\biggl((a, b] \biggr) \quad &= \quad \sum_{k=1}^{n} \tau_g\biggl((a_k, b_k] \biggr) \, + \, \tau_g\biggl(\cup_{k=n+1}^{\infty} (a_k, b_k] \biggr) \\ &\geq \quad \sum_{k=1}^{n} \tau_g\biggl((a_k, b_k] \biggr) \, + \, \tau_g\biggl(\cup \cup_{k=n+1}^{K} (a_k, b_k] \biggr) \\ &= \quad \sum_{k=1}^{n} \tau_g\biggl((a_k, b_k] \biggr) \, + \, \sum_{k=n+1}^{K} \tau_g\biggl((a_k, b_k] \biggr). \end{aligned}$$

We know that

$$\lim_{K} \sum_{k=n+1}^{K} \tau_g\biggl((a_k, b_k] \biggr) \quad = \quad 0.$$

Thus, letting $K \to \infty$, we find for all $n > N$, that

$$\tau_g\Big((a, b] \Big) \quad \geq \quad \sum_{k=1}^{n} \tau_g\Big((a_k, b_k] \Big).$$

However, the sequence of partial sums above converges. We have then the inequality

$$\tau_g\Big((a, b] \Big) \quad \geq \quad \sum_{k=1}^{\infty} \tau_g\Big((a_k, b_k] \Big).$$

Combining the two inequalities, we have that our extension τ_g is a pseudo-measure. ∎

Comment 13.3.1. *It is worthwhile to summarize what we have accomplished at this point. We know now that the premeasure τ_g defined by the nondecreasing and right continuous map g on the algebra of sets, $\mathcal{A}$, generated by the collection $\mathcal{U}$ consisting of the empty set, finite intervals like $(a, b]$ and unbounded intervals of the form $(-\infty, b]$ and (a, ∞) when defined to be additive on $\mathcal{A}$ generates an interesting outer measure μ_b^*. We have also proven that the extension τ_g becomes a pseudo-measure on $\mathcal{A}$. Thus,*

- *(i): The sets A in $\mathcal{A}$ are in the σ - algebra of sets that satisfy the Caratheodory condition using μ_g^* which we denote by $\mathcal{M}_g$. We denote the resulting measure by μ_g.*
- *(ii): We know μ_g is regular and complete.*
- *(iii): We know that $\mu_g(A) = \tau_g(A)$ for all A in $\mathcal{A}$.*
- *(iv): Since any open set can be written as a countable disjoint union of open intervals, this means any open set is in $\mathcal{M}_g$ because $\mathcal{M}_g$ contains open intervals as they are in $\mathcal{A}$ and the σ - algebra $\mathcal{M}_g$ is closed under countable disjoint unions. This also tells us that the Borel σ - algebra is contained in $\mathcal{M}_g$.*

We can also prove that μ_g^* is an outer measure. Since open sets are μ_g^* measurable, by Theorem 12.2.2, it follows that μ_g^* is a metric outer measure.

Comment 13.3.2. *The measures μ_g induced by the outer measures μ_g^* are called Lebesgue - Stieljes measures . Since open sets are measurable here, these measures are also called Borel measures .*

Comment 13.3.3. *So for a given nondecreasing right continuous g, we can construct a Lebesgue - Stieljes measure satisfying*

$$\mu_g\Big((a, b] \Big) \quad = \quad g(b) \; - \; g(a).$$

So what about the open interval (a, b)? We know that

$$(a, b) \quad = \quad \bigcup_n (a, b \; - \; \frac{1}{n}].$$

Then

$$\begin{aligned} \mu_g\Big((a, b) \Big) \quad &= \quad \lim_n \, g(b \; - \; \frac{1}{n}) \; - \; g(a) \\ &= \quad g(b^-) \; - \; g(a). \end{aligned}$$

What about the singleton $\{b\}$? We know

$$\{b\} \quad = \quad \bigcap_n (b \, - \, \frac{1}{n}, b]$$

and so

$$\begin{aligned} \mu_g\Big(\{b\}\Big) \quad &= \quad \lim_n \, g(b) \, - \, g(b \, - \, \frac{1}{n}) \\ &= \quad g(b) \, - \, g(b^-). \end{aligned}$$

Note this tells us that the Lebesgue - Stieljes measure of a singleton need not be 0. *However, at any point* b *where* g *is continuous, this measure will be zero. Since our* g *can have at most a countable number of discontinuities, we see there are only a countable number of singleton sets whose measure is non - zero.*

13.4 Homework

Exercise 13.4.1. *A family* $\mathcal{A}$ *of subsets of the set* X *is an* **algebra** *if*

(i): $\emptyset$, X *are in* $\mathcal{A}$.

(ii): $E \in \mathcal{A}$ *implies* $E^C \in \mathcal{A}$.

(iii): if $\{A_1, \ldots, A_n\}$ *is a finite collection of sets in* $\mathcal{A}$, *then their union is in* $\mathcal{A}$.

Further, the mapping τ *is sometimes called a pseudo-measure on the algebra* $\mathcal{A}$ *if* $\tau : \mathcal{A} \rightarrow [0, \infty]$ *and*

(i): $\tau(\emptyset) = 0$.

(ii): If (A_i) *is a countable collection of disjoint sets in* $\mathcal{A}$ *whose union is also in* $\mathcal{A}$ *(note this is not always true because* $\mathcal{A}$ *is not a* σ *- algebra), then*

$$\tau(\cup_i \, A_i) \quad = \quad \sum_i \, \tau(A_i).$$

Now we get to the exercise:

(a): Let $\mathcal{U}$ *be the family of subsets of* $\Re$ *of the form* $(a, b]$, $(-\infty, b]$, (a, ∞) *and* $(-\infty, \infty)$ *as well as* $\emptyset$. *Prove* $\mathcal{F}$, *the collection of all finite unions of sets from* $\mathcal{U}$ *is an algebra of subsets of* $\Re$.

(b): Prove τ *equal to the usual length of an interval is a pseudo-measure on* $\mathcal{F}$.

(c): Let g *be any monotone increasing function on* $\Re$ *which is continuous from the right. This means*

$$\begin{aligned} \lim_{h \rightarrow 0^+} g(x + h) \;\; &\textit{exists} \;, \; \forall \, x, \\ \lim_{x \rightarrow -\infty} g(x) \;\; &\textit{exists}, \\ \lim_{x \rightarrow \infty} g(x) \;\; &\textit{exists}. \end{aligned}$$

where the last two limits could be $-\infty$ and ∞ respectively. Define the mapping τ_g on $\mathcal{U}$ by

$$\begin{aligned}
\tau_g\Big((a,b]\Big) &= g(b) - g(a), \\
\tau_g\Big((-\infty,b)\Big) &= g(b) - \lim_{x\to-\infty} g(x), \\
\tau_g\Big((a,\infty)\Big) &= \lim_{x\to\infty} g(x) - g(a), \\
\tau_g\Big((-\infty,\infty)\Big) &= \lim_{x\to\infty} g(x) - \lim_{x\to-\infty} g(x).
\end{aligned}$$

and extend τ_g to $\mathcal{F}$ as usual. Prove that τ_g is a pseudo-measure on $\mathcal{F}$.

(d): τ_g can then be used to define an outer measure μ_g^* as usual. There is then an associated σ - algebra of μ_g^* measurable sets of $\Re$, $\mathcal{M}_g$, and μ_g^* restricted to $\mathcal{M}_g$ is a measure, μ_g.

We now prove $\mathcal{F}$ is contained in $\mathcal{M}_g$. Here is the hint for any set I from $\mathcal{F}$. Compare this problem to Example 12.4.1 and Example 12.4.2 which are almost identical in spirit (although the g here is more general) even though they are couched in terms of pre-measures instead of pseudo-measures.

Hint. *Let T be any subset of $\Re$. Let $\epsilon > 0$ be given. Then there is a cover (A_n) of sets from the algebra $\mathcal{F}$ so that*

$$\sum_n \tau_g(A_n) \leq \mu_g^*(T) + \epsilon.$$

Now $I \cap T \subseteq \cup_n (A_n \cap I)$ and $I^C \cap T \subseteq \cup_n (A_n \cap I^C)$. So

$$\begin{aligned}
\mu_g^*(T \cap I) &\leq \sum_n \tau_g(A_n \cap I), \\
\mu_g^*(T \cap I^C) &\leq \sum_n \tau_g(A_n \cap I^C).
\end{aligned}$$

Combining, and using the additivity of τ_g, we see

$$\mu_g^*(T \cap I) + \mu_g^*(T \cap I^C) \leq \sum_n \tau_g(A_n) \leq \mu_g^*(T) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have shown I satisfies the Caratheodory condition and so in μ_g^* measurable.

Once you have shown these things, we know the Borel σ - algebra $\mathcal{B}$ is contained in $\mathcal{M}_g$! Measures constructed this way are called Borel - Stieljes measures on $\Re$ when we restrict them to $\mathcal{B}$. If we use the full σ - algebra, we call them Lebesgue - Stieljes measures.

Exercise 13.4.2. *Let h be our Cantor function*

$$h(x) = (x + \Psi(x))/2.$$

From the previous exercise, we know τ_h defines a Borel - Stieljes measure. Determine if τ_h is absolutely continuous with respect to the Borel measure on $\Re$ (Borel measure is just Lebesgue measure restricted to $\mathcal{B}$.

Chapter 14

Modes Of Convergence

There are many ways a sequence of functions in a measure space can converge. In this chapter, we will explore some of them and the relationships between them.

There are several types of convergence here:

(i): Convergence pointwise,

(ii): Convergence uniformly,

(iii): Convergence almost uniformly,

(iv): Convergence in measure,

(v): Convergence in $\mathcal{L}_p$ norm for $1 \leq p < \infty$,

(vi): Convergence in $\mathcal{L}_\infty$ norm.

We will explore each in turn. We have already discussed the p norm convergence in Chapter 11 so there is no need to go over those ideas again. However, some of the other types of convergence in the list above are probably not familiar to you. Pointwise and pointwise a.e. convergence have certainly been mentioned before, but let's make a formal definition so it is easy to compare it to other types of convergence later.

Definition 14.0.1. Convergence Pointwise and Pointwise a.e.
Let $(X, \mathcal{S})$ be a measurable space. Let (f_n) be a sequence of extended real valued measurable functions: i.e. $(f_n) \subseteq M(X, \mathcal{S})$. Let $f : X \to \overline{\Re}$ be a function. Then, we say f_n converges pointwise to f on X if $\lim_n f_n(x) = f(x)$ for all x in X. Note that this type of convergence does not involve a measure although it does use the standard metric, $||$ on $\Re$. We can write this as

$$f_n \to f \ [ptws].$$

If there is a measure μ on $\mathcal{S}$, we can also say the sequence converges almost everywhere if $\mu(\{x \mid f_n(x) \not\to f(x)\}) = 0$. We would write this as

$$f_n \to f \ [ptws\, a.e.].$$

Next, you have probably already seen uniform convergence in the context of advanced calculus. We can define it nicely in a measure space also.

Definition 14.0.2. Convergence Uniformly
Let $(X, \mathcal{S})$ be a measurable space. Let (f_n) be a sequence of real valued measurable functions: i.e. $(f_n) \subseteq M(X, \mathcal{S})$. Let $f : X \to \overline{\Re}$ be a function. Then, we say f_n converges uniformly to f on X if for any $\epsilon > 0$, there is a positive integer N (depending on the choice of ϵ so that if $n > N$, then $\mid f_n(x) - f(x) \mid < \epsilon$ for all x in X. We can write this as

$$f_n \to f \ [unif].$$

However, if we are in a measure space, we can relax the idea of uniform convergence of the whole space by taking advantage of the underlying measure.

Definition 14.0.3. Almost Uniform Convergence
Let $(X, \mathcal{S}\mu)$ be a measure space. Let $(f_n) \subseteq M(X, \mathcal{S}, \mu)$ be a sequence of functions which are finite a.e. Let $f : X \to \overline{\Re}$ be a function. We say f_n converges almost uniformly to f on X if for any $\epsilon > 0$, there is a measurable set E such that $\mu(E^C) < \epsilon$ and (f_n) converges uniformly to f on E. We write this as

$$f_n \to f \ [a.u.].$$

Finally, we can talk about a brand new idea: convergence using only measure itself.

Definition 14.0.4. Convergence In Measure

Let $(X, \mathcal{S}\mu)$ be a measure space. Let $(f_n) \subseteq M(X, \mathcal{S}, \mu)$ be a sequence of functions which are finite a.e. Let $f : X \to \overline{\Re}$ be a function. Let E be a measurable set. We say f_n converges in measure to f on E if for any pair (ϵ, δ) of positive numbers, there exists a positive integer N (depending on ϵ and δ) so that if $n > N$, then

$$\mu(\{x \mid \mid f_n(x) - f(x) \mid \geq \delta\}) < \epsilon.$$

We write this as

$$f_n \to f \; [meas \, on \, E].$$

If E is all of X, we would just write

$$f_n \to f \; [meas].$$

14.1 Subsequence Extraction

In some cases, when a sequence of functions converges in one way, it is possible to prove that there is at least one subsequence that converges in a different manner. We will now make this idea precise.

Definition 14.1.1. Cauchy Sequences In Measure

Let $(X, \mathcal{S}, \mu)$ be a measure space and (f_n) be a sequence of extended real valued measurable functions. We say (f_n) is Cauchy in Measure if for all $\alpha > 0$ and $\epsilon > 0$, there is a positive integer N so that

$$\mu\Big(|f_n(x) - f_m(x)| \geq \alpha\Big) < \epsilon, \; \forall \, n, \, m > N.$$

We can prove a kind of completeness result next.

Theorem 14.1.1. Cauchy In Measure Implies A Convergent Subsequence

Let $(X, \mathcal{S}, \mu)$ be a measure space and (f_n) be a sequence of extended real valued measurable functions which is Cauchy in Measure. Then there is a subsequence (f_n^1) and an extended real valued measurable function f such that $f_n^1 \to f$ [a.e.], $f_n^1 \to f$ [a.u.] and $f_n^1 \to f$ [meas].

Proof. *For each pair of indices n and m, there is a measurable set E_{nm} on which the definition of the difference $f_n - f_m$ is not defined. Hence, the set*

$$E = \bigcup_n \bigcup_m E_{nm}$$

is measurable and on E^C, all differences are well defined. We do not know the sets E_{nm} have measure 0 here as the members of the sequence do not have to be summable or essentially bounded.

Now, let's get started with the proof.

(Step 1): let $\alpha_1 = 1/2$ and $\epsilon_1 = 1/2$ also. Then, (f_n) Cauchy in Measure implies

$$\exists \, N_1 \ni n, \, m > N_1 \;\; \Rightarrow \;\; \mu\Big(|f_n(x) - f_m(x)| \geq 1/2\Big) < 1/2.$$

Let

$$g_1 \;=\; f_{N_1+1}.$$

(Step 2): let $\alpha_2 = 1/2^2$ *and* $\epsilon_1 = 1/2^2$ *also. Then,* (f_n) *Cauchy in Measure again implies there is an* $N_2 > N_1$ *so that*

$$n,\, m \;>\; N_2 \;\;\Rightarrow\;\; \mu\Bigg(|f_n(x) \;-\; f_m(x)| \;\geq\; 1/4\Bigg) \;<\; 1/4.$$

Let

$$g_2 \;=\; f_{N_2+1}.$$

It is then clear by our construction that

$$\mu\Bigg(|g_2(x) \;-\; g_1(x)| \;\geq\; 1/2\Bigg) \;<\; 1/2.$$

(Step 3): let $\alpha_3 = 1/2^3$ *and* $\epsilon_1 = 1/2^3$ *also. Then,* (f_n) *Cauchy in Measure again implies there is an* $N_3 > N_2$ *so that*

$$n,\, m \;>\; N_3 \;\;\Rightarrow\;\; \mu\Bigg(|f_n(x) \;-\; f_m(x)| \;\geq\; 1/8\Bigg) \;<\; 1/8.$$

Let

$$g_3 \;=\; f_{N_3+1}.$$

It follows by construction that

$$\mu\Bigg(|g_3(x) \;-\; g_2(x)| \;\geq\; 1/4\Bigg) \;<\; 1/4.$$

Continuing this process by induction, we find a subsequence (g_n) *of the original sequence* (f_n) *so that for all* $k \geq 1$,

$$\mu\Bigg(|g_{k+1}(x) \;-\; g_k(x)| \;\geq\; 1/2^k\Bigg) \;<\; 1/2^k.$$

Define the sets

$$E_j \;=\; \Bigg(|g_{j+1}(x) \;-\; g_j(x)| \;\geq\; 1/2^j\Bigg)$$

and

$$F_k \;=\; \bigcup_{j=k}^{\infty} E_j.$$

Note if $x \in F_k^C$,

$$|g_{j+1}(x) \;-\; g_j(x)| \;\;<\;\; 1/2^j$$

for any index $j \geq k$. Each set F_k is then measurable and they form an increasing sequence. Let's get a bound on $\mu(F_k)$. First, if A and B are measurable sets, then

$$\mu\Bigl(A \cup B\Bigr) \;=\; \mu\Bigl(A \cup B^C\Bigr) : + \mu\Bigl(A \cap B\Bigr) \,+\, \mu\Bigl(A^C \cup B\Bigr)$$

But adding in $\mu\Bigl(A \cap B\Bigr)$ simply makes the sum larger. We see

$$\begin{aligned} \mu\Bigl(A \cup B\Bigr) &\leq \mu\Bigl(A \cup B^C\Bigr) : + \mu\Bigl(A \cap B\Bigr) \,+\, \mu\Bigl(A \cap B\Bigr) \,+\, \mu\Bigl(A^C \cup B\Bigr) \\ &= \mu(A) \,+\, \mu(B). \end{aligned}$$

This result then extends easily to finite unions. Thus, if (A_n) is a sequence of measurable sets, then by the sub additive result above,

$$\mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) \;\leq\; \sum_{i=1}^{n} \mu(A_i).$$

Hence, the sets $\cup_{i=1}^{n} A_i$ form an increasing sequence and we clearly have

$$\mu\Bigl(\bigcup_{i=1}^{\infty} A_i\Bigr) \;=\; \lim_n \, \mu\Bigl(\bigcup_{i=1}^{n} A_i\Bigr) \;\leq\; \sum_{i=1}^{\infty} \mu(A_i).$$

We can apply this idea to the increasing sequence (F_k) to obtain

$$\begin{aligned} \mu(F_k) &\leq \sum_{j=k}^{\infty} \mu(E_j) \\ &< \sum_{j=k}^{\infty} 1/2^j \;=\; 1/2^{k-1}. \end{aligned}$$

Now, for any $i > j$, we have

$$|g_i(x) \,-\, g_j(x)| \;\leq\; \sum_{\ell=j}^{i-1} |g_{\ell+1} \,-\, g_\ell|.$$

Choosing the indices i and j so that $i > j \geq k$, we then find if $x \notin F_k$, that

$$|g_{\ell+1}(x) \,-\, g_\ell(x)| \;<\; 1/2^\ell.$$

Hence, for these indices,

$$\begin{aligned} |g_i(x) \,-\, g_j(x)| &\leq \sum_{\ell=j}^{i-1} |g_{\ell+1} \,-\, g_\ell| \\ &< \sum_{\ell=j}^{i-1} 1/2^\ell \;=\; \sum_{\ell=j}^{\infty} 1/2^\ell \;=\; 1/2^{j-1}. \end{aligned}$$

We conclude that if $x \in F_k^C$ and $i > j \geq k$ we have

$$|g_i(x) - g_j(x)| \leq 1/2j - 1. \qquad (*)$$

Now let $F = \cap_k F_k$. The F is measurable and $\mu(F) = \lim_k \mu(F_k) = 0$. Let x be in F^C. By De Morgan's Laws, $x \in \cup_k F_K^C$ which implies x is in some F_k^C. Call this k^*. Then given $\epsilon > 0$, choose J so that $1/2^{J-1} < \epsilon$. Then, by Equation $*$, if $i > j \geq J \geq k^*$,

$$|g_i(x) - g_j(x)| \leq 1/2^{j-1} < 1/2^{J-1} < \epsilon.$$

Thus, the sequence $g_k(x)$ is a Cauchy sequence of real numbers for each x in F^C. Hence, $\lim_k g_k(x)$ exists for such x. Defining f by

$$f(x) = \begin{cases} \lim_k g_k(x), & x \in F^C \\ 0, & x \in F, \end{cases}$$

we see f is measurable and it is the pointwise limit a.e. of the subsequence (g_k). This completes the proof of the first claim. To see that (g_k) converges in measure to f, look again at Equation $*$:

$$|g_i(x) - g_j(x)| \leq 1/2j - 1, \ \forall\, i > j \geq k, \ \forall\, x \in F_k^C.$$

Now let $i \to \infty$ and use the continuity of the absolute value function to obtain

$$|f(x) - g_j(x)| \leq 1/2j - 1, \ \forall\, j \geq k, \ \forall\, x \in F_k^C. \qquad (**)$$

Equation $**$ says that (g_k) converges to f uniformly on F_k^C. Further, recall $\mu(F_k) < 1/2^{k-1}$. Note given any $\delta > 0$, there is an integer k^* so that $1/2^{k^*-1} < \delta$ and g_k converges uniformly on $F_{k^*}^C$. We therefore conclude that (g_k) converges almost uniformly to f as well.

To show the last claim, given an arbitrary $\alpha > 0$ and $\epsilon > 0$, choose a positive integer k^* so that

$$\mu(F_k^*) < 1/2^{k^*-1} < \min(\alpha, \epsilon).$$

Then, by Equation $**$, we have

$$\Big(|f(x) - g_j(x)| \geq \alpha\Big) \subseteq \Big(|f(x) - g_j(x)| > 1/2^{k^*-1}\Big).$$

Then, again by Equation $**$, we have

$$\begin{aligned} &\subseteq \Big(|f(x) - g_j(x)| > 1/2^{k^*-1}\Big) \\ &\subseteq F_{k^*}. \end{aligned}$$

Combining, we have

$$\mu\Big(|f(x) - g_j(x)| \geq \alpha\Big) \leq \mu\Big(F_{k^*}\Big) < 1/2^{k^*-1} < \epsilon.$$

This shows that (g_k) converges to f in measure. ■

The result above allows us to prove that Cauchy in Measure implies there is a function which the Cauchy sequence actually converges to.

Theorem 14.1.2. Cauchy In Measure Implies Completeness
Let $(X, \mathcal{S}, \mu)$ be a measure space and (f_n) be a sequence of extended real valued measurable functions which is Cauchy in Measure. Then there is an extended real valued measurable function f such that $f_n \rightarrow f$ [meas] and the limit function f is determined uniquely a.e.

Proof. *By Theorem 14.1.1, there is a subsequence (f_n^1) and a real valued function measurable function f so that $f_n^1 \rightarrow f$ [meas]. Let $\alpha > 0$ be given. If $|f(x) - f_n(x)| \geq \alpha$, then given any f_n^1 in the subsequence, we have*

$$\alpha \leq |f(x) - f_n(x)| \quad \leq \quad |f(x) - f_n^1(x)| + |f_n(x) - f_n^1(x)|.$$

Note, just as in the previous proof, there is a measurable set E where all additions and subtractions of functions are well-defined. Now, let $\beta = |f(x) - f_n^1(x)|$ and $\gamma = |f_n(x) - f_n^1(x)|$. The equation above thus says

$$\beta + \gamma \quad \leq \quad \alpha$$

Since β and γ are non negative and both are less than or equal to α, we can think about this inequality in a different way. If there was equality

$$\beta^* + \gamma^* \quad = \quad \alpha$$

with both β^* and γ^* not zero, then we could let $t = \beta^*/\alpha$ and we could say $\beta^* = t\,\alpha$ and $\gamma^* = (1-t)\,\alpha$ as $\gamma^* = \alpha - \beta^*$. Now imagine β and γ being larger α. Then, β and γ would have to be bigger than or equal to the values $\beta^* = t\,\alpha$ and $\gamma^* = (1-t)\,\alpha$ for some t in $(0, 1)$. Similar arguments work for the cases of $\beta = 0$ and $\gamma = 0$ which will correspond to the cases of $t = 0$ and $t = 1$. Hence, we can say that if $|f(x) - f_n(x)| \geq \alpha$, then there is some $t \in [0, 1]$ so that

$$\begin{aligned} |f(x) - f_n^1(x)| &\geq t\,\alpha, \\ |f_n(x) - f_n^1(x)| &\geq (1-t)\,\alpha. \end{aligned}$$

The following reasoning is a bit involved, so bear with us. First, if x is a value where $|f(x) - f_n(x)| \geq \alpha$, we must have that $|f(x) - f_n^1(x)| \geq t\,\alpha$ (call this Condition I) **and** *$|f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha$ (call this Condition II).*
Case (i): if $0 \leq t \leq 1/2$, then since an x which satisfies Condition I must also satisfy Condition II, we see for these values of t, we have

$$\begin{aligned} \{x \mid |f(x) - f_n^1(x)| \geq t\,\alpha\} &\subseteq \{x \mid |f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \\ &\subseteq \{x \mid |f_n(x) - f_n^1(x)| \geq 1/2\,\alpha\}. \end{aligned}$$

Hence, for $0 \leq t \leq 1/2$, we conclude

$$\{x \mid |f(x) - f_n^1(x)| \geq t\,\alpha\} \bigcup \{x \mid |f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \quad \subseteq \quad \{x \mid |f_n(x) - f_n^1(x)| \geq 1/2\,\alpha\}.$$

A similar argument shows that if $1/2 \leq t \leq 1$, any x satisfying Condition II must satisfy Condition I. Hence, for these t,

$$\begin{aligned} \{x \mid |f(x) - f_n^1(x)| \geq t\,\alpha\} \quad &\bigcup \quad \{x \mid |f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \\ &\subseteq \quad \{x \mid |f(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \\ &\subseteq \quad \{x \mid |f(x) - f_n^1(x)| \geq 1/2\,\alpha\}. \end{aligned}$$

Combining these results, we find

$$\begin{aligned} &\bigcup_{0 \leq t \leq 1} \left(\{x \mid |f(x) - f_n^1(x)| \geq t\,\alpha\} \bigcup \{x \mid |f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \right) \\ &\qquad \subseteq \{x \mid |f_n(x) - f_n^1(x)| \geq 1/2\,\alpha\} \bigcup \{x \mid |f(x) - f_n^1(x)| \geq 1/2\,\alpha\} \qquad . \end{aligned}$$

Finally, from the triangle inequality,

$$|f(x) - f_n(x)| \quad \leq \quad |f(x) - f_n^1(x)| \; + \; |f_n^1(x) - f_n(x)|,$$

and so, we have

$$\begin{aligned} \{x \mid |f(x) - f_n(x)| \geq \; \alpha\} \quad &\subseteq \quad \bigcup_{0 \leq t \leq 1} \left(\{x \mid |f(x) - f_n^1(x)| \geq t\,\alpha\} \bigcup \{x \mid |f_n(x) - f_n^1(x)| \geq (1-t)\,\alpha\} \right) \\ &\subseteq \quad \{x \mid |f_n(x) - f_n^1(x)| \geq 1/2\,\alpha\} \bigcup \{x \mid |f(x) - f_n^1(x)| \geq 1/2\,\alpha\}. \end{aligned}$$

Next, pick an arbitrary $\epsilon > 0$. Since $f_n^1 \to f$ [meas], there is a positive integer N_1 so that

$$\mu\Big(|f(x) - f_n^1(x)| \; \geq \; \alpha/2 \Big) \quad < \quad \epsilon/2, \; \forall\, n^1 \; > \; N_1.$$

where n^1 denotes the index of the function f_n^1. Further, since (f_n) is Cauchy in measure, there is a positive integer N_2 so that

$$\mu\Big(|f_n(x) - f_n^1(x)| \; \geq \; \alpha/2 \Big) \quad < \quad \epsilon/2, \; \forall\, n, \, n^1 \; > \; N_2.$$

So if n^1 is larger than $N = \max(N_1, N_2)$, we have

$$\mu\Big(|f(x) - f_n(x)| \; \geq \; \alpha/2 \Big) \quad < \quad \epsilon, \; \forall\, n \; > \; N.$$

This shows $f_n \to f$ [meas] as desired.

To show the uniqueness a.e. of f, assume there is another function g so that $f_n \rightarrow g$ [meas]. Then, by arguments similar to ones we have already used, we find

$$\{x \mid |f(x) - g(x)| \geq \alpha\} \quad \subseteq \quad \{x \mid |f_n(x) - f(x)| \geq 1/2\,\alpha\}.$$

Then, mutatis mutandi, we obtain

$$\begin{aligned} \mu\Big(\{x \mid |f(x) - g(x)| \geq \alpha\}\Big) &\leq \mu\Big(\{x \mid |f_n(x) - f(x)| \geq 1/2\,\alpha\}\Big) \\ &+ \mu\Big(\{x \mid |f_n(x) - g(x)| \geq 1/2\,\alpha\}\Big) \\ &< \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we see for any $\alpha > 0$,

$$\mu\Big(\{x \mid |f(x) - g(x)| \geq \alpha\}\Big) \quad = \quad 0.$$

However, we know

$$\mu\Big(\{x \mid |f(x) - g(x)| > 0\}\Big) \quad = \quad \bigcup_n \Big(\{x \mid |f(x) - g(x)| \geq 1/n\}\Big),$$

which immediately tells us that

$$\mu\Big(\{x \mid |f(x) - g(x)| > 0\}\Big) \quad = \quad 0.$$

This says $f = g$ a.e. and we are done. ■

Theorem 14.1.3. p-Norm Convergence Implies Convergence in Measure

Assume $1 \leq p < \infty$. Let (f_n) be a sequence in $\mathcal{L}_p(X, \mathcal{S}, \mu)$ and let $f \in \mathcal{L}_p(X, \mathcal{S}, \mu)$ so that $f_n \rightarrow f$ $[p - norm]$. Then $f_n \rightarrow f$ [meas] which is Cauchy in Measure.

Proof. *Let $\alpha > 0$ be given and let*

$$E_n(alpha) \quad = \quad \{x \mid |f_n(x) \; - \; f(x)| \; \geq \; \alpha\}.$$

Then, given $\epsilon > 0$, there is a positive integer N so that

$$\int |f_n - f|^p \, d\mu \quad < \quad \alpha^p \, \epsilon, \;\; \forall \, n \; > \; N.$$

Thus,

$$\int_{E_n(\alpha)} |f_n - f|^p \, d\mu \leq \int |f_n - f|^p \, d\mu \quad < \quad \alpha^p \, \epsilon, \;\; \forall \, n \; > \; N.$$

But on $E_n(\alpha)$, the integrand in the first term is bigger than or equal to α^p. We obtain

$$\alpha^p \, \mu(E_n(\alpha)) \quad < \quad \alpha^p \, \epsilon, \; \forall \, n \, > \, N.$$

Canceling the α^p term, we have $\mu(E_n(\alpha)) < \epsilon$, for all $n > N$. This implies $f_n \to f$ [meas]. ■

Comment 14.1.1. *Let's assess what we have learned so far. We have shown*

(i):

$$f_n \;\to\; f \,[p - norm] \quad \Rightarrow \quad f_n \;\to\; f \,[meas]$$

by Theorem 14.1.3.

(ii): It is a straightforward exercise to show

$$f_n \;\to\; f \,[meas] \quad \Rightarrow \quad (f_n) \;\; \textit{Cauchy In Measure} \; .$$

Then,

$$(f_n) \;\; \textit{Cauchy In Measure} \quad \Rightarrow \quad \exists \, (f_n^1) \, \subseteq (f_n) \, \ni \, f_n^1 \;\to\; f \,[a.e.]$$

by Theorem 14.1.1. Note, we proved the existence of such a subsequence already in the proof of the completeness of $\mathcal{L}_p$ as discussed in Theorem 11.1.10.

(iii): Finally, we can also apply Theorem 14.1.1 to infer

$$f_n \;\to\; f \,[meas] \quad \Rightarrow \quad \exists \, (f_n^1) \, \subseteq (f_n) \, \ni \, f_n^1 \;\to\; f \,[a.u.]$$

Theorem 14.1.4. Almost Uniform Convergence Implies Convergence In Measure
Let $(X, \mathcal{S})$ be a measurable space. Let (f_n) be a sequence of real valued measurable functions: i.e. $(f_n) \subseteq M(X, \mathcal{S})$. Let $f : X \to \overline{\Re}$ be measurable. Then

$$f_n \;\to\; f \,[a.u.] \quad \Rightarrow \quad f_n \;\to\; f \,[meas].$$

Proof. *If f_n converges to f a.u., given arbitrary $\epsilon > 0$, there is a measurable set E_ϵ so that $\mu(E_\epsilon) < \epsilon$ and f_n converges uniformly on E_ϵ^C. Now let $\alpha > 0$ be chosen. Then, there is a positive integer N_α so that*

$$|f_n(x) \; - \; f(x)| \quad < \quad \epsilon, \; \forall \, n \, > \, N_\alpha, \; \forall \, x \, \in \, E_\epsilon^C.$$

Hence, if $n > N_\alpha$ and x satisfies $|f_n(x) \; - \; f(x)| \geq \alpha$, we must have that $x \in E_\epsilon$. We conclude

$$\Big(|f_n(x) \; - \; f(x)| \; \geq \; \alpha \Big) \quad \subseteq \quad E_\epsilon, \; \forall \, n \, > \, N_\alpha.$$

This implies immediately that

$$\mu\Big(|f_n(x) \; - \; f(x)| \, \geq \, \alpha\Big) \quad \leq \quad \mu(E_\epsilon) \; < \; \epsilon, \; \forall \, n \; > \; N_\alpha.$$

This proves $f_n \; \rightarrow \; f$ *[meas].* ■

Comment 14.1.2. *We have now shown*

$$f_n \; \rightarrow \; f \, [p - norm] \quad \Rightarrow \quad f_n \; \rightarrow \; f \, [meas]$$

by Theorem 14.1.3. This then implies by Theorem 14.1.1

$$\exists \, (f_n^1) \; \subseteq (f_n) \; \ni \; f_n^1 \; \rightarrow \; f \, [a.u.]$$

14.2 Egoroff's Theorem

A famous theorem tells us how pointwise a.e. convergence can be phrased "almost" like uniform convergence. This is **Egoroff's Theorem**.

Theorem 14.2.1. Egoroff's Theorem
Let $(X, \mathcal{S}, \mu)$ *be a measure space with* $\mu(X) < \infty$*. Let* f *be an extended real valued function which is measurable. Also, let* (f_n) *be a sequence of functions in* $M(X, \mathcal{S})$ *such that* $f_n \; \rightarrow \; f$ *[a.e.]. Then,* $f_n \; \rightarrow \; f$ *[a.u.] and* $f_n \; \rightarrow \; f$ *[meas].*

Proof. *From previous arguments, the way we handle converge a.e. is now quite familiar. Also, we know how we deal with the measurable set on which addition of the function* f_n *are not well defined. Hence, we may assume without loss of generality that the convergence here is on all of* X *and that addition is defined on all of* X*. With that said, let*

$$E_{nk} \quad = \quad \bigcup_{k=n}^{\infty} \Big(|f_k(x) - f(x)| \, \geq \, 1/m\Big).$$

Note that each E_{nk} *is measurable and* $E_{n+1,k} \subseteq E_{nk}$ *so that this is an decreasing sequence of sets in the index* n*. Given* x *in* X*, we have* $f_n \rightarrow f(x)$*. Hence, for* $\epsilon = 1/m$*, there is a positive integer* $N(x, \epsilon)$ *so that*

$$|f_n(x) \; - \; f(x)| \quad < \quad \epsilon \; = \; 1/m, \; \forall \, n \; > \; N(x, \epsilon).$$

Thus,

$$\Big(|f_n(x) \; - \; f(x)| \, \geq \, 1/m\Big) \quad = \quad \emptyset, \; \forall \, n \; > \; N(x, \epsilon). \tag{$*$}$$

Now consider $F_m \; = \; \bigcap_{n=1}^{\infty} E_{nm}$*. If* $x \, \in \, F_m$*, then* x *is in* E_{nm} *for all* n*. In particular, letting* $n^* = N(x, \epsilon) + 1$*, we have* $x \in E_{n^* \, m}$*. Looking at how we defined* $E_{n^* \, m}$*, we see this implies that there is a positive integer* $k' > n^*$*, so that* $|f_k'(x) - f(x)| \, \geq \, 1/m$*. However, by Equation* $*$*, this set is empty.*

This contradiction means our original assumption that F_m was non empty is wrong. Hence, $F_m = \emptyset$. Now, since $\mu(X) < \infty$, $\mu(E_{1\,m}$ is finite also. Hence, by Lemma 10.1.2,

$$0 \;=\; \mu(F_m) \;\;=\;\; \lim_n \mu(E_{n\,m}.$$

This implies that given $\delta > 0$, there is a positive integer N_m so that

$$\mu(E_{n\,m} \;\;<\;\; \delta/2^m, \;\forall\, m \;>\; N_m,$$

since $\lim_m \mu(E_{nm} = 0$. For each integer m, choose a positive integer $n_m > N_m$. We can arrange for these integers to be increasing; i.e., $n_m < n_{m+1}$. Then,

$$\mu\Big(E_{n_m\,m}\Big) \;\;<\;\; \delta/2^m$$

and letting

$$E_\delta \;\;=\;\; \bigcup_{m=1}^{\infty} E_{n_m\,m},$$

we have

$$\mu(E_\delta) \;\;\leq\;\; \sum_{m=1}^{\infty} \delta/2^m \;<\; \delta.$$

Finally, note

$$E_\delta^C \;\;=\;\; \left(\bigcup_{m=1}^{\infty} E_{n_m\,m}\right)^C \;=\; \bigcap_{m=1}^{\infty} E_{n_m\,m}^C.$$

Next, note

$$\begin{aligned} E_{n_m\,m}^C \;\;&=\;\; \left(\bigcup_{k=n_m}^{\infty}\Big(|f_k(x) \;-\; f(x)| \;\geq\; 1/m\Big)\right)^C \\ &=\;\; \bigcap_{k=n_m}^{\infty}\Big(|f_k(x) \;-\; f(x)| \;\geq\; 1/m\Big)^C \\ &=\;\; \bigcap_{k=n_m}^{\infty}\Big(|f_k(x) \;-\; f(x)| \;<\; 1/m\Big). \end{aligned}$$

Thus, since $x \in E_\delta^C$ means x is in $E_{n_m\,m}^C$ for all m, the above says $|f_k(x) - f(x)| < 1/m$ for all $k > n_m$. Therefore, given $\epsilon > 0$, pick a positive integer M so that $1/M < \epsilon$. Then, for all x in E_δ^C, we have

$$|f_k(x) \;-\; f(x)| \;\;<\;\; 1/M \;<\; \epsilon, \;\forall\, k \;\geq n_M.$$

This says f_n converges uniformly to f on E_δ^C with $\mu(E_\delta) < \delta$. Hence, we have shown $f_n \;\rightarrow\; f$ [a.u.]

Finally, if $f_n \;\rightarrow\; f$ [a.u.], by Theorem 14.1.4, we have $f_n \;\rightarrow\; f$ [meas] also. ■

Next, let's see what we can do with domination by a p-summable function.

Theorem 14.2.2. Pointwise a.e. Convergence Plus Domination Implies p-Norm Convergence

Let $1 \leq p < \infty$ and $(X, \mathcal{S}, \mu)$ be a measure space. Let f be an extended real valued function which is measurable. Also, let (f_n) be a sequence of functions in $\mathcal{L}_p(X, \mathcal{S})$ such that $f_n \to f[a.e.]$. Assume there is a dominator function g which is p-summable; i.e. $|f_n(x)| \leq g(x)$ a.e. Then, if $f_n \to f$ [a.e.], f is p-summable and $f_n \to f$ $[p - norm]$.

Proof. *Since $|f_n(x)| \leq g(x)$ a.e., we have immediately that $|f| \leq g$ a.e. since $f_n \to f$ [a.e.]. Thus, $|f|^p \leq g^p$ and we know f is in $\mathcal{L}_p(X, \mathcal{S})$. Since all the functions here are p-summable, the set where all additions is not defined has measure zero. So, we can assume without loss of generality that this set has been incorporated into the set on which convergence fails. Hence, we can say*

$$|f_n(x) - f(x)| \leq |f_n(x)| + |f(x)| \leq 2\,g(x), \text{ a.e.}$$

So,

$$|f_n(x) - f(x)|^p \leq 2^p\,|g(x)|^p, \text{ a.e.}$$

By assumption, g is p-summable, so we have $2^p\,g^p$ is in $\mathcal{L}_1(X, \mathcal{S})$. Applying Lebesgue's Dominated Convergence Theorem, we find

$$\lim_n \int |f_n(x) - f(x)|^p\,d\mu = \int \lim_n |f_n(x) - f(x)|^p\,d\mu = 0.$$

Thus, $f_n \to f$ $[p - norm]$. ∎

14.3 Vitali Convergence Theorem

This important theorem is one that gives us more technical tools to characterize p-norm convergence for a sequence of functions. We need a certain amount of technical infrastructure to pull this off; so bear with us as we establish a series of lemmatta.

Lemma 14.3.1. p-Summable Functions Have p-Norm Arbitrarily Small Off a Set

Let $1 \leq p < \infty$ and $(X, \mathcal{S}, \mu)$ be a measure space. Let f be in $\mathcal{L}_p(X, \mathcal{S})$. Then given $\epsilon > 0$, there is a measurable set E_ϵ so that $\mu(E_\epsilon) < \infty$ and if $F \subseteq E_\epsilon^C$ is measurable, then $\|\, f\, I_F \,\|_p < \epsilon$.

Proof. *Let $E_n = (|f_n(x)| \geq 1/n)$. Note $E_n \in \mathcal{S}$ and the sequence (E_n) is increasing and $\cup_n E_n = X$. Let $f_n = f I_{E_n}$. It is straightforward to verify that $f_n \uparrow f$ as $f_n \leq f_{n+1}$ for all n. Further, $|f_n|^p \leq |f|^p$; hence, by the Dominated Convergence Theorem,*

$$\lim_n \int |f_n|^p\,d\mu = \int \lim_n |f_n|^p\,d\mu = \int |f|^p\,d\mu < \infty.$$

The definition of f_n and E_n then implies

$$\mu(E_n)/n^p \quad \leq \quad \int_{E_n} |f|^p \, d\mu \leq \int |f|^p \, d\mu < \infty.$$

This tells us $\mu(E_n) < \infty$ for all n.
Now choose $\epsilon > 0$ arbitrarily. Then there is a positive integer N so that

$$\int |f|^p \, d\mu - \int |f_n|^p \, d\mu \quad < \quad \epsilon^p, \, \forall \, n \, > \, N.$$

Thus, since $f_n = f I_{E_n}$, we can say

$$\int_{E_n} |f|^p \, d\mu + \int_{E_n^C} |f|^p \, d\mu - \int_{E_n} |f|^p \, d\mu \quad < \quad \epsilon^p, \, \forall \, n \, > \, N.$$

or

$$\int_{E_n^C} |f|^p \, d\mu \quad < \quad \epsilon^p, \, \forall \, n \, > \, N.$$

So choose $E_\epsilon = E_{N+1}$ and we have

$$\int_{E_\epsilon^C} |f|^p \, d\mu \quad < \quad \epsilon^p.$$

which implies the desired result. ■

Lemma 14.3.2. p-Summable Inequality

Let $1 \leq p < \infty$ and $(X, \mathcal{S}, \mu)$ be a measure space. Let (f_n) be a sequence of functions in $\mathcal{L}_p(X, \mathcal{S})$. Define β_n on $\mathcal{S}$ by

$$\beta_n(E) \quad = \quad \| \, f_n \, I_E \, \|_p, \, \forall \, E.$$

Then,

$$|\beta_n(E) \; - \; \beta_m(E)| \quad \leq \quad \| \, f_n - f_m \, \|_p, \, \forall \, E, \, \forall \, n, \, m.$$

Proof. *By the backwards triangle inequality, for any measurable E,*

$$\| \, f_n \; - \; f_m \, \|_p \quad \geq \quad | \, \| \, f_n \, I_E \, \|_p \; - \; \| \, f_m \, I_E \, \|_p \, | \; = \; |\beta_n(E) \; - \; \beta_m(E)|.$$

■

Lemma 14.3.3. p-Summable Cauchy Sequence Condition I

Let $1 \leq p < \infty$ and $(X, \mathcal{S}, \mu)$ be a measure space. Let (f_n) be a Cauchy Sequence in $\mathcal{L}_p(X, \mathcal{S})$. Define β_n on $\mathcal{S}$ as done in Lemma 14.3.2. Then, there is a positive integer N and a measurable set E_ϵ of finite measure, so that if F is a measurable subset of E_ϵ, then $\beta_n(E) \; < \; \epsilon$ for all $n > N$.

Proof. *Since (f_n) is a Cauchy sequence in p-norm, there is a function f in $\mathcal{L}_p(X, \mathcal{S})$ so that $f_n \rightarrow f\ [p - norm]$. By Lemma 14.3.1, given $\epsilon > 0$, there is a measurable set E_ϵ with finite measure so that*

$$\int_{E_\epsilon^C} |f|^p \, d\mu \quad < \quad (\epsilon/2)^p.$$

Now given a measurable F contained in E_ϵ^C, recalling the meaning of $\beta_n(F)$ as described in Lemma 14.3.2, we can write

$$\begin{aligned} \beta_n(F) \quad &\leq \quad \| \, f_n \, \|_p \leq \| \, f_n \, I_{E_\epsilon^C} \, \|_p \\ &\leq \quad \| \, (f_n - f) \, I_{E_\epsilon^C} \, \|_p \; + \; \| \, f \, I_{E_\epsilon^C} \, \|_p \\ &< \quad \epsilon/2 + \; \| \, (f_n - f) \, I_{E_\epsilon^C} \, \|_p \; . \end{aligned}$$

Since $f_n \rightarrow f\ [p - norm]$, there is a positive integer N so that if $n > N$,

$$\| \, (f_n - f) \, I_{E_\epsilon^C} \, \|_p \quad < \quad \epsilon/2.$$

This shows $\beta_n(F) < \epsilon$ when $n > N$ as desired. ■

Lemma 14.3.4. Continuity Of The Integral

Let $(X, \mathcal{S}, \mu)$ be a measure space and f be a summable function. Then for all $\epsilon > 0$ there is a $\delta > 0$, so that

$$|\int_E f \, d\mu| \quad < \quad \epsilon, \forall \, E \; \in \; \mathcal{S}, \;\; with \;\; \mu(E) \; < \; \delta.$$

Proof. *Define the measure γ on $\mathcal{S}$ by $\gamma(E) = \int_E |f| \, d\mu$. Note, by Comment 10.4.1, we know that γ is absolutely continuous with respect to μ. Now assume the proposition is false. Then, there is an $\epsilon > 0$ so that for all choices of $\delta > 0$, we have a measurable set E_δ for which $\mu(E_\delta) < \delta$ and $|\int_{E_\delta} f \, d\mu|/geq\epsilon$. In particular, for the sequence $\delta_n = 1/2^n$, we have a sequence of sets E_n with $\mu(E_n) < 1/2^n$ and*

$$|\int_{E_n} f \, d\mu|/geq\epsilon.$$

Let

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

Then,

$$\mu(G) \quad \leq \quad \mu(G_n) \; \leq \; \sum_{k=n}^{\infty} \mu(E_k) \; < \; \sum_{k=n}^{\infty} 1/2^k \; = \; 1/2^{n-1}.$$

This implies $mu(G) = 0$ and thus, since γ is absolutely continuous with respect to μ, $\gamma(G) = 0$ also. We also know the sequence G_n is decreasing and so $\gamma(G_n) \rightarrow \gamma(G) = 0$. Finally, since

$$\gamma(G_n) \quad \geq \quad \gamma(E_n) \; \geq \; |\int_E f \, d\mu| \; \geq \; \epsilon,$$

we have $\gamma(G) = \lim_n \gamma(G_n) \geq \epsilon$ *as well. This is impossible. Hence, our assumption that the proposition is false is wrong.* ■

Lemma 14.3.5. p-Summable Cauchy Sequence Condition II

Let $1 \leq p < \infty$ *and* $(X, \mathcal{S}, \mu)$ *be a measure space. Let* (f_n) *be a Cauchy Sequence in* $\mathcal{L}_p(X, \mathcal{S})$*. Define* β_n *on* $\mathcal{S}$ *as done in Lemma 14.3.2. Then, given* $\epsilon > 0$*, there is a* $\delta > 0$ *and a positive integer* N *so that if* $n > N$*, then*

$$\beta_n(E) \;<\; \epsilon, \forall\, E \in \mathcal{S}, \;\; \textit{with} \;\; \mu(E) < \delta.$$

Proof. *Since* $\mathcal{L}_p(X, \mathcal{S}, \mu)$ *is complete, there is a p-summable function* f *so that* $f_n \rightarrow f$ $[p-norm]$*. Then, by Lemma 14.3.4, given an* $\epsilon > 0$*, there is a* $\delta > 0$*, so that*

$$\int_E |f|^p \, d\mu \;<\; (\epsilon/2)^p, \;\; \textit{if} \;\; \mu(E) \;<\; \delta.$$

Hence, using the convenience mapping $\beta_n(E)$ *previously defined in Lemma 14.3.2, we see*

$$\begin{aligned} \beta_n(E) &= \| f_n \, I_E \|_p = \| (f - f_n) \, I_E \|_p \;+\; \| f \, I_E \|_p \\ &\leq \| (f - f_n) \, I_E \|_p \;+ \epsilon/2 \end{aligned}$$

when $\mu(E) < \delta$*. Finally, since* $f_n \rightarrow f$ $[p-norm]$*, there is a positive integer* N *so that if* $n > N$*, then* $\| (f - f_n) \, I_E \|_p < \epsilon/2$*. Combining, we have* $\beta_n(E) < \epsilon$ *if* $n > N$ *for* $\mu(E) < \delta$*.* ■

Theorem 14.3.6. Vitali Convergence Theorem

Let $1 \leq p < \infty$ *and* $(X, \mathcal{S}, \mu)$ *be a measure space. Let* (f_n) *be a sequence of functions in* $\mathcal{L}_p(X, \mathcal{S})$*. Then,* $f_n \rightarrow f$ $[p-norm]$ *if and only if the following three conditions hold.*

(i):

$$f_n \rightarrow f \;\; [meas]$$

(ii):

$$\forall\, \epsilon > 0, \, \exists\, N, \, \exists\, E_\epsilon \in \mathcal{S}, \mu(E_\epsilon) < \infty, \ni \; F \subseteq E_\epsilon^C, \, F \in \mathcal{S} \;\; \Rightarrow \;\; \int_F |f_n|^p \, d\mu \;<\; \epsilon^p, \;\forall\, n \;>\; N.$$

(iii):

$$\forall\, \epsilon > 0, \, \exists\, \delta > 0, \, \exists\, N \ni \; E \in \mathcal{S}, \, \mu(E) \;<\; \delta \;\; \Rightarrow \;\; \int_E |f_n|^p \, d\mu \;<\; \epsilon^p, \;\forall\, n \;>\; N.$$

Proof.
$\Rightarrow$*: If* $f_n \rightarrow f$ $[p-norm]$*, then by Theorem 14.1.3,* $f_n \rightarrow f$ $[meas]$ *which shows (i) holds. Then, since* $f_n \rightarrow f$ $[p-norm]$*,* (f_n) *a Cauchy sequence. Thus, by Lemma 14.3.1, condition (ii) holds. Finally, since* (f_n) *is a Cauchy sequence, by Lemma 14.3.5, condition (iii) holds.*

⇒: Now assume conditions (i), (ii) and (iii) hold. Let $\epsilon > 0$ be given. From condition (ii), we see there is a measurable set E_ϵ of finite measure and a positive integer N_1 so that

$$\int_{E_\epsilon^C} |f_n|^p \, d\mu \quad < \quad (\epsilon/4)^p$$

if $n > N_1$. Thus, for indices n and m larger than N_1, we have

$$\begin{aligned} \| f_n - f_m \|_p &= \| (f_n - f_m)\, I_{E_\epsilon} + (f_n - f_m)\, I_{E_\epsilon^C} \|_p \\ &\leq \| (f_n - f_m)\, I_{E_\epsilon} \|_p + \| (f_n - f_m)\, I_{E_\epsilon^C} \|_p \\ &\leq \| (f_n - f_m)\, I_{E_\epsilon} \|_p + \| f_n\, I_{E_\epsilon^C} \|_p + \| f_m\, I_{E_\epsilon^C} \|_p \\ &< \| (f_n - f_m)\, I_{E_\epsilon} \|_p + \epsilon/2. \end{aligned}$$

We conclude

$$\| f_n - f_m \|_p \quad < \quad \| f_m\, I_{E_\epsilon} \|_p + \epsilon/2, \, \forall\, n,\, m \, > \, N_1. \tag{$*$}$$

Now let $\beta = \mu(E_\epsilon$ and set

$$\alpha \quad = \quad \frac{\epsilon}{4\, \beta^{1/p}}$$

and define the sets H_{nm} by

$$H_{nm} \quad = \quad \{x \mid |f_n(x) - f_m(x)| \, \geq \, \alpha\}.$$

Apply condition (ii) for our given ϵ now. Thus, there is a $\delta(\epsilon)$ and a positive integer N_2 so that

$$\int_E |f_n|^p \, d\mu \quad < \quad (\epsilon/8)^p, \; n \, > \, N - 2, \quad \text{when} \;\; \mu(E) \, < \, \delta(\epsilon). \tag{$**$}$$

Since $f_n \rightarrow f$ [meas], (f_n) is Cauchy in measure. Hence, there is a positive integer N_3 so that

$$\mu(H_{nm} \quad < \quad \delta(\epsilon), \, \forall\, n,\, m \, > \, N_3. \tag{$* * *$}$$

Finally, using the Minkowski Inequality, we have

$$\begin{aligned} \| (f_n - f_m)\, I_{E_\epsilon} \|_p &= \| (f_n - f_m)\, I_{E_\epsilon \setminus H_{nm}} + (f_n - f_m)\, I_{H_{nm}} \|_p \\ &\leq \| (f_n - f_m)\, I_{E_\epsilon \setminus H_{nm}} \|_p + \| f_n\, I_{H_{nm}} \|_p + \| f_m\, I_{H_{nm}} \|_p \end{aligned}$$

Now let $N = \max\{N_1, N_2, N_3\}$*. Then, if* n *and* m *exceed* N*, we have Equation* $*$*, Equation* $**$ *and Equation* $***$ *all hold. This implies*

$$\begin{aligned} \| (f_n - f_m)\, I_{E_\epsilon} \|_p &\leq \Big(\alpha^p\, \mu(E_\epsilon \setminus H_{nm})\Big)^{1/p} + \epsilon/8 + \epsilon/8 \\ \& \leq \alpha \Big(\mu(E_\epsilon)\Big)^{1/p} + \epsilon/4 \\ &= \alpha\, \beta^{1/p} + \epsilon/4 = \epsilon/(4\,\beta^{1/p})\, \beta^{1/p} + \epsilon/4 \\ &= \epsilon/2. \end{aligned}$$

From Equation $*$*, we have for these indices* n *and* m*,*

$$\begin{aligned} \| f_n - f_m \|_p &< \| f_m\, I_{E_\epsilon} \|_p + \epsilon/2 \\ &< \epsilon. \end{aligned}$$

Thus, (f_n) *is a Cauchy sequence in p-norm. Since* $\mathcal{L}_p(X, \mathcal{S}, \mu)$ *is complete, there is a function* g *so that* $f_n \to g\,[p-norm]$*. So by Theorem 14.1.3,* $f_n \to g$ $[meas]$*. It is then straightforward to show that* $f = g$ *a.e. This tells us* f *and* g *belong to the same equivalence class of* $\mathcal{L}_p(X, \mathcal{S}, \mu)$*.* ■

14.4 Summary

We can summarize the results of this chapter as follows. If the measure of X is infinite, we have many one way implications.

Theorem 14.4.1. Convergence Relationships On General Measurable Space
Let $(X, \mathcal{S}, \mu)$ *be a measure space. Let* f *and* (f_n) *be in* $M(X, \mathcal{S})$*. Then, we know the following implications:*

(i):

$$\left.\begin{array}{l} f_n \to f\ [p-norm] \\ f_n \to f\ [unif] \\ f_n \to f\ [a.u.] \end{array}\right\} \Rightarrow f_n \to f\ [meas].$$

(ii):

$$\left.\begin{array}{l} f_n \to f\ [unif] \\ f_n \to f\ [a.u.] \end{array}\right\} \Rightarrow f_n \to f\ [a.e.].$$

(iii):

$$f_n \to f\ [unif] \Rightarrow f_n \to f\ [a.u.].$$

If we know the measure is finite, we can say more.

Theorem 14.4.2. Convergence Relationships On Finite Measure Space

Let $(X, \mathcal{S}, \mu)$ be a measure space with $\mu(X) < \infty$. Let f and (f_n) be in $M(X, \mathcal{S})$. Then, we know the following implications:

(i):

$$\left.\begin{array}{l} f_n \rightarrow f \ [p - norm] \\ f_n \rightarrow f \ [unif] \\ f_n \rightarrow f \ [a.u.] \\ f_n \rightarrow f \ [a.e.] \end{array}\right\} \Rightarrow \quad f_n \rightarrow f \ [meas].$$

(ii):

$$\left.\begin{array}{l} f_n \rightarrow f \ [unif] \\ f_n \rightarrow f \ [a.u.] \end{array}\right\} \Rightarrow \quad f_n \rightarrow f \ [a.e.].$$

(iii):

$$f_n \rightarrow f \ [unif] \quad \Rightarrow \quad \left\{\begin{array}{l} f_n \rightarrow f \ [a.u.] \\ f_n \rightarrow f \ [p - norm] \end{array}\right.$$

(iv):

$$f_n \rightarrow f \ [a.e.]. \quad \Rightarrow \quad f_n \rightarrow f \ [a.u.].$$

Next, if we can dominate the sequence by an $\mathcal{L}_p$ function, we can say even more.

Theorem 14.4.3. Convergence Relationships With p-Domination

Let $(X, \mathcal{S}, \mu)$ be a measure space. Let f and (f_n) be in $M(X, \mathcal{S})$. Assume there is a $g \in L_p$ so that $|f_n| \leq g$. Then, we know the following implications:

(i):

$$\left.\begin{array}{l} f_n \rightarrow f \ [p-norm] \\ f_n \rightarrow f \ [unif] \\ f_n \rightarrow f \ [a.u.] \\ f_n \rightarrow f \ [a.e.] \end{array}\right\} \Rightarrow f_n \rightarrow f \ [meas].$$

(ii):

$$\left.\begin{array}{l} f_n \rightarrow f \ [unif] \\ f_n \rightarrow f \ [a.u.] \end{array}\right\} \Rightarrow f_n \rightarrow f \ [a.e.].$$

(iii):

$$f_n \rightarrow f \ [unif] \Rightarrow \left\{\begin{array}{l} f_n \rightarrow f \ [a.u.] \\ f_n \rightarrow f \ [p-norm] \end{array}\right.$$

(iv):

$$f_n \rightarrow f \ [a.e.]. \Rightarrow f_n \rightarrow f \ [a.u.].$$

(v):

$$\left.\begin{array}{l} f_n \rightarrow f \ [a.e.] \\ f_n \rightarrow f \ [a.u.] \end{array}\right\} \Rightarrow f_n \rightarrow f \ [p-norm].$$

(vi):

$$f_n \rightarrow f \ [meas]. \Rightarrow f_n \rightarrow f \ [p-norm].$$

There are circumstances where we can be sure we can extract a subsequence that converges in some fashion.

Theorem 14.4.4. Convergent Subsequences Exist

Let $(X, \mathcal{S}, \mu)$ be a measure space. It doesn't matter whether or not $\mu(X)$ is finite. Let f and (f_n) be in $M(X, \mathcal{S})$. Then, we know the following implications:

(i):

$$\left.\begin{array}{l} f_n \rightarrow f \ [p-norm] \\ f_n \rightarrow f \ [meas] \end{array}\right\} \Rightarrow \exists \ subsequence \ f_n^1 \rightarrow f \ [a.u.].$$

(ii):

$$\left.\begin{array}{l} f_n \rightarrow f \ [p-norm] \\ f_n \rightarrow f \ [meas] \end{array}\right\} \Rightarrow \exists \ subsequence \ f_n^1 \rightarrow f \ [a.e.].$$

Further, the same implications hold if we know there is a $g \in L_p$ so that $|f_n| \leq g$.

14.5 Homework

Exercise 14.5.1. *Characterize convergence in measure when the measure in counting measure.*

Exercise 14.5.2. *Let $(X, \mathcal{S}\mu)$ be a measure space. Let $(f_n), (g_n) \subseteq M(X, \mathcal{S}, \mu)$ be sequences of functions which are finite a.e. Let $f, g : X \rightarrow \overline{\Re}$ be functions. Prove if $f_n \rightarrow f[meas\, on\, E]$ and $g_n \rightarrow g[meas\, on\, E]$, then $(f_n + g_n) \rightarrow (f+g)[meas\ on\ E]$.*

Exercise 14.5.3. *Let $(X, \mathcal{S}\mu)$ be a measure space with $\mu(X) < \infty$. Let $(f_n), (g_n) \subseteq M(X, \mathcal{S}, \mu)$ be sequences of functions which are finite a.e Let $f, g : X \rightarrow \overline{\Re}$ be functions. Prove if $f_n \rightarrow f[meas on E]$ and $g_n \rightarrow g[meas\, on\, E]$, then $(f_n g_n) \rightarrow (f g)[meas\, on\, E]$. Hint: first consider the case that $f_n \rightarrow 0[meas\, on\, E]$ and $g_n \rightarrow 0[meas\ on\ E]$.*

Exercise 14.5.4. *Let $(X, \mathcal{S}, \mu)$ be a measure space. Let $(f_n) \subseteq M(X, \mathcal{S}, \mu)$ be a sequence of functions which are finite a.e. Let $f : X \rightarrow \overline{\Re}$ be a function. Prove if $f_n \rightarrow f[a.u.]$, then $f_n \rightarrow f[ptws\ a.e.]$ and $f_n \rightarrow f[meas]$.*

Exercise 14.5.5. *Let $(X, \mathcal{S}, \mu)$ be a finite measure space. for any pair of measurable functions f and g, define*

$$d(f,g) = \int \frac{|\ f-g\ |}{1+|\ f-g\ |}\, d\mu.$$

(i): Prove $M(X, \mathcal{S}, \mu)$ is a semi-metric space.

(ii): Prove if (f_n) is a sequence of measurable functions and f is another measurable function, then $f_n \rightarrow f[meas]$ if and only if $d(f_n, f) \rightarrow 0$.

Hint: You don't need any high power theorems here. First, let $\phi(t) = t/(1+t)$ so that $d(f,g) = \int \phi(|f-g| d\mu$. Then try this:
$(\Rightarrow)$: We assume $f_n \rightarrow f[meas]$. Then, given any pair of positive numbers (δ, ϵ), we have there is an N so that if $n > N$, we have

$$\mu(|f_n(x) - f(x)| \geq \delta) < \epsilon/2.$$

Let E_δ denote the set above. Now for such $n > N$, note

$$d(f_n, f) \;=\; \int_{E_\delta} \phi(|f_n - f|)\, d\mu \;+\; \int_{E_\delta^C} \phi(|f_n - f|)\, d\mu.$$

Since ϕ is increasing, we see that on E_δ^C, $\phi(|f_n(x) - f(x)|) < \phi(\delta)$. Thus, you should be able to show that if $n > N$, we have

$$d(f_n, f) \;<\; \mu(E_\delta) \;+\; \phi(\delta)\, \mu(X) \;=\; \epsilon/2 \;+\; \phi(\delta)\, \mu(X).$$

Then a suitable choice of δ does the job.
($\Leftarrow$): If we know $d(f_n, f)$ goes to zero, break the integral up the same way into a piece on E_δ and E_δ^C. This tells us right away that given $\epsilon > 0$, there is an N so that $n > N$ implies

$$\phi(\delta)\, \mu(E_\delta) \;<\; \epsilon.$$

This gives us the result with a little manipulation.

Exercise 14.5.6. *Let $(\Re, \mathcal{M}, \mu)$ denote the measure space consisting of the Lebesgue measurable sets $\mathcal{M}$ and Lebesgue measure μ. Let the sequence (f_n) of measurable functions be defined by*

$$f_n \quad = \quad n\, I_{[1/n, 2/n]}.$$

Prove $f_n \to 0$ on all $\Re$, $f_n \to 0$ [meas] but $f_n \not\to 0$ [$p - norm$] for $1 \le p < \infty$.

Chapter 15

Decomposition Of Measures

We now examine the structure of a charge λ on a σ - algebra $\mathcal{S}$. For convenience, let's recall that a charge is a mapping on $\mathcal{S}$ to $\Re$ which assigns the value 0 to $\emptyset$ and which is countably additive. We need some beginning definitional material before we go further.

15.1 Basic Decomposition Results

Definition 15.1.1. Positive and Negative Sets For a Charge

Let λ be a charge on $(X, \mathcal{S})$. We say $P \in \mathcal{S}$ is a positive set with respect to λ if

$$\lambda(E \cap P) \geq 0, \forall E \in \mathcal{S}.$$

Further, we say $N \in \mathcal{S}$ is a negative set with respect to λ if

$$\lambda(E \cap N) \leq 0, \forall E \in \mathcal{S}.$$

Finally, $M \in \mathcal{S}$ is a null set with respect to λ is

$$\lambda(E \cap M) = 0, \forall E \in \mathcal{S}.$$

Definition 15.1.2. The Positive and Negative Parts Of a Charge

Let λ be a charge on $(X, \mathcal{S})$. Define the mapping λ^+ on $\mathcal{S}$ by

$$\lambda^+(E) \quad = \quad \sup \{ \lambda(A) \mid A \in \mathcal{S},\ A \subseteq E \}.$$

Also, define the mapping λ^- on $\mathcal{S}$ by

$$\lambda^-(E) \quad = \quad -\inf \{ \lambda(A) \mid A \in \mathcal{S},\ A \subseteq E \}.$$

Theorem 15.1.1. The Jordan Decomposition Of A Charge

Let λ be a charge on $(X, \mathcal{S})$. Then, λ^+ and λ^- are finite measures on $\mathcal{S}$ and $\lambda = \lambda^+ - \lambda^-$. The pair (λ^+, λ^-) is called the Jordan Decomposition of λ.

Proof. *Let's look at λ^+ first. Given any measurable E, since $\emptyset$ is contained in E, by the definition of λ^+, we must have $\lambda^+(E) \geq \lambda(\emptyset) = 0$. Hence, λ^+ is non negative.*
Next, if A and B are measurable and disjoint, By definition of λ^+, for $C_1 \subseteq A$ $C_2 \subseteq B$, we must have

$$\begin{aligned} \lambda^+(A \cup B) &\geq \lambda(C_1 \cup C_2) \\ &= \lambda(C_1) + \lambda(C_2). \end{aligned}$$

This says $\lambda^+(A \cup B) - \lambda(C_2)$ is an upper bound for the set of numbers $\{\lambda(C_1)\}$. Hence, by definition of $\lambda^+(A)$, we have

$$\begin{aligned} \lambda^+(A \cup B) &\geq \lambda(C_1 \cup C_2) \\ &= \lambda^+(A) + \lambda(C_2). \end{aligned}$$

A similar argument then shows that $\{\lambda(C_2)\}$ is bounded above by $\lambda^+(A \cup B) - \lambda^+(A)$. Thus, we have

$$\begin{aligned} \lambda^+(A \cup B) &\geq \lambda(C_1 \cup C_2) \\ &= \lambda^+(A) + \lambda^+(B). \end{aligned}$$

On the other hand, if $C \subseteq A \cup B$, then we have

$$\begin{aligned} \lambda(C) &= \lambda(C \cap A \cup C \cap B) \\ &\leq \lambda^+(A) + \lambda^+(B). \end{aligned}$$

This immediately implies that

$$\lambda^+(A \cup B) \quad \leq \quad \lambda^+(A) + \lambda^+(B).$$

Thus, it is clear λ^+ is additive on finite disjoint unions.
We now address the question of the finiteness of λ^+. To see λ^+ is finite, assume that it is not. So there is some set E with $\lambda^+(E) = \infty$. Hence, by definition, there is a measurable set A_1 so that $\lambda(A_1) > 1$.

Thus, by additivity of λ^+, we have

$$\lambda^+(A_1) + \lambda^+(E \setminus A_1) = \lambda^+(E) = \infty.$$

Thus, at least one of of $\lambda^+(A_1)$ and $\lambda^+(E \setminus A_1)$ is also ∞. Pick one such a set and call it B_1. Thus, $\lambda^+(B_1) = \infty$. Let's do one more step. Since $\lambda^+(B_1) = \infty$, there is a measurable set A_2 inside it so that $\lambda(A_2) > 2$. Then,

$$\lambda^+(A_2) + \lambda^+(B_1 \setminus A_2) = \lambda^+(B_1) = \infty.$$

Thus, at least one of of $\lambda^+(A_2)$ and $\lambda^+(B_1 \setminus A_2)$ is also ∞. Pick one such a set and call it B_2. Thus, we have $\lambda^+(B_2) = \infty$. You should be able to see how we construct the two sequences (A_n) and (B_n). When we are done, we know $A_n \subseteq B_{n-1}$, $\lambda(A_n) > n$ and $\lambda(B_n) = \infty$ for all n.

Now, if for an infinite number of indices n_k, $B_{n_k} = B_{n_k - 1} \setminus A_{n_k}$, what happens? It is easiest to see with an example. Suppose $B_5 = B_4 \setminus A_5$ and $B_8 = B_7 \setminus A_8$. By the way we construct these sets, we see A_6 does not intersect A_5. Hence, $A_7 \cap A_5 = \emptyset$ also. Finally, we have $A_8 \cap A_5 = \emptyset$ too. Hence, extrapolating from this simple example, we can infer that the sequence $(A_{n_k}$ is disjoint. By the countable additivity of λ, we then have

$$\lambda\Bigl(\bigcup_k A_{n_k}\Bigr) = \sum_k \lambda(A_{n_k} > \sum_k n_k = \infty.$$

But λ is finite on all members of $\mathcal{S}$. This is therefore a contradiction.

Another possibility is that there is an index N so that if $n > N$, the choice is always that of $B_n = A_n$. In this case, we have

$$E \supseteq A_{N+1} \supseteq A_{N+2} \ldots$$

Since λ is finite and additive,

$$\lambda(A_{N+j-1} \setminus A_{N+j}) = \lambda(A_{N+j-1}) - \lambda(A_{N+j})$$

for $j > 2$ since all the λ values are finite. We now follow the construction given in the proof of the second part of Lemma 10.1.2 to finish our argument. Construct the sequence of sets (E_n) by

$$\begin{aligned} E_1 &= \emptyset \\ E_2 &= A_{N+1} \setminus A_{N+2} \\ E_3 &= A_{N+1} \setminus A_{N+3} \\ \vdots &\ \vdots\ \vdots \\ E_n &= A_{N+1} \setminus A_{N+n-1}. \end{aligned}$$

Then (E_n) is an increasing sequence of sets which are disjoint and so $\lambda(\cup_n E_n) = \lim_n \lambda(E_n)$. Since $\lambda(A_{N+1})$ is finite, we then know that $\lambda(E_n) = \lambda(A_{N+1}) - \lambda(A_{N+n})$. Hence, $\lambda(\cup_n E_n) = \lambda(A_{N+1}) -$

$\lim_n \lambda(A_{N+n})$. *Next, note by De Morgan's Laws,*

$$\begin{aligned}
\lambda\Big(\cup_n E_n\Big) &= \lambda\Big(\bigcup_n A_{N+1} \cap A^C_{N+n}\Big) \\
&= \lambda\Big(A_{N+1} \bigcap \cup_n A^C_{N+n}\Big) \\
&= \lambda\Big(A_{N+1} \bigcap \Big(\cap_n A_{N+n}\Big)^C\Big) \\
&= \lambda\Big(A_{N+1} \setminus \Big(\cap_n A_{N+n}\Big)\Big).
\end{aligned}$$

Thus, since $\lambda(A_{N+1})$ is finite and $\cap_n A_{N+n} \subseteq A_{N+1}$, it follows that

$$\lambda(\cup_n E_n) = \lambda(A_{N+1}) - \lambda(\cap_n A_{N+n}).$$

Combining these results, we have

$$\lambda(A_{N+1}) - \lim_n \lambda(A_{N+n}) = \lambda(A_{N+1}) - \lambda(\cap_n A_{N+n}).$$

Canceling $\lambda(A_{N+1})$ from both sides, we find

$$\lambda(\cap_n A_{N+n}) = \lim_n \lambda(A_{N+n}) \geq \lim_n N + n = \infty.$$

We again find a set $\cap_n A_{N+n}$ with λ value ∞ inside E. However, λ is always finite. Thus, in this case also, we arrive at a contradiction.

We conclude at this point that if $\lambda^+(E) = \infty$, we force λ to become infinite for some subsets. Since that is not possible, we have shown λ^+ is finite. Since $\lambda^- = (-\lambda)^+$, we have established that λ^- is finite also. Next, given the relationship between λ^+ and λ^-, it is enough to prove λ^+ is a measure to complete this proof.

It is enough to prove that λ^+ is countably additive. Let (E_n) be a countable sequence of measurable sets and let E be their union. If $A \subseteq E$, then $A = \cup_n A \cap E_n$ and so

$$\begin{aligned}
\lambda(A) &= \sum_n \lambda(A \cap E_n) \\
&\leq \sum_n \lambda^+(E_n),
\end{aligned}$$

by the definition of λ^+. Since this holds for all such subsets A, we conclude $\sum_n \lambda^+(E_n)$ is an upper bound for the collection of all such $\lambda(A)$. Hence, by the definition of a supremum, we have $\lambda^+(E) \leq \sum_n \lambda^+(E_n)$.

To show the reverse, note $\lambda^+(E)$ is finite by the arguments in the first part of this proof. Now, pick $\epsilon > 0$. Then, by the Supremum Tolerance Lemma, there is a sequence (A_n) of measurable sets, each $A_n \subseteq E_n$ so that

$$\lambda^+(E_n) - \epsilon/2^n < \lambda(A_n) \leq \lambda^+(E_n).$$

Let $A = \cup_n A_n$. Then $A \subseteq E$ and so we have $\lambda(A) \leq \lambda^+(E)$. Hence,

$$\begin{aligned} \sum_n \lambda^+(E_n) &< \sum_n \Big(\lambda(A_n) + \epsilon/2^n \Big) \\ &< \sum_n \lambda(A_n) + \epsilon \end{aligned}$$

since the second term is a standard geometric series. Next, since (A_n) is a disjoint sequence, the countable additivity of λ gives

$$\sum_n \lambda^+(E_n) < \lambda\Big(\cup_n A_n \Big) + \epsilon.$$

But $A = \cup_n A_n$ and since this holds for all $\epsilon > 0$, we can conclude

$$\sum_n \lambda^+(E_n) < \lambda(A) \leq \lambda^+(E).$$

Combining these inequalities, we see λ^+ is countably additive and hence is a measure. ■

Comment 15.1.1. *If we had allowed the charge in Definition 10.0.2 to be extended real valued; i.e. take on the values of ∞ and $-\infty$, what would happen? First, note by applying the arguments in the first part of the proof above, we can say if $\lambda^+(E) = \infty$, $\lambda(E) = \infty$ and similarly, if $\lambda^-(E) = \infty$, $\lambda(E) = -\infty$. Conversely, note by definition of λ^+, if $\lambda(E) = \infty$, then $\lambda^+(E) = \infty$ also and if $\lambda(E) = -\infty$, then $\lambda^-(E) = \infty$. So if $\lambda^+(E) = \infty$, what about $\lambda^-(E)$? If $\lambda^-(E) = \infty$, that would force $\lambda(E) = -\infty$ contradicting the value it already has. Hence $\lambda^-(E)$ is finite. Next, given any measurable set F, what about $\lambda^-(F)$? There are several cases. First, if $F \subseteq E$, then*

$$\lambda^-(E) \& = \lambda^-(F) + \lambda^-(F \setminus E).$$

Since $\lambda^- \geq 0$, if $\lambda^-(F) = \infty$, we get $\lambda^-(E)$, which is finite, is also infinite. Hence, this can not happen. Second, if F and E are disjoint, with $\lambda^-(F) = \infty$, we find

$$\lambda^-(E \cup F) = \lambda^-(E) + \lambda^-(F).$$

The right hand side is ∞ and so since $\lambda^-(E \cup F)$ is infinite, this forces $\lambda(E \cup F) = -\infty$. But since λ is additive on disjoint sets, this leads to the undefined expression

$$\begin{aligned} \lambda(E \cup F) &= \lambda(E) + \lambda(F) \\ -\infty &= (\infty) + (-\infty). \end{aligned}$$

This is not possible because by assumption, λ takes on a well defined value in $\overline{\Re}$ for all measurable subsets. Thus, we conclude if there is a measurable set E so that $\lambda^+(E)$ is infinite, then λ^- will be finite everywhere. The converse is also true: if $\lambda^-(E)$ is infinite, then λ^+ will be finite everywhere. Thus, we can conclude if λ is extended real valued, only one of λ^+ or λ^- can take on ∞ values.

Now we show that any charge λ has associated Positive and Negative sets.

Theorem 15.1.2. The Hahn Decomposition Associated With A Charge
Let λ be a charge on $(X, \mathcal{S})$. Then, there is a positive set P and a negative set N so that $X = P \cup N$ and $P \cap N = \emptyset$. The pair (P, N) is called a Hahn Decomposition associated with the charge λ.

Proof. *Since λ^+ is finite, first by the Supremum Tolerance Lemma there are measurable sets A_n so that*

$$\lambda(A_n) \;>\; \lambda^+(X) - 1/2^n \tag{$*$}$$

for all n. Hence, by the Jordan Decomposition of λ, we can say

$$\lambda^-(A_n) \;=\; \lambda^+(A_n) - \lambda(A_n) \le \lambda^+(X) - \lambda(A_n) < 1/2^n, \tag{$**$}$$

by Equation $*$.

Next, note if E is measurable and in $X \setminus A_n$, then

$$\lambda(E) + \lambda(A_n) \;=\; \lambda(E \cup A_n) \le \lambda^+(X)$$

by the definition of λ^+. Hence,

$$\lambda(E) \;\le\; \lambda^+(X) - \lambda(A_n) < 1/2^n,$$

again by Equation $*$. This immediately implies

$$\lambda^+(X \setminus A_n) \;\le\; 1/2^n. \tag{$***$}$$

Now, with these preliminaries out of the way, let

$$A \;=\; \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \;=\; \limsup(A_n).$$

Then, a simple application of DeMorgan's Laws gives

$$X \setminus A \;=\; \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n^C \;=\; \liminf(A_n).$$

Thus, since λ^- is a measure,

$$\begin{aligned} \lambda^-\Big(A\Big) &\le \lambda^-\Big(\bigcup_{n=k}^{\infty} A_n\Big) \\ &\le \sum_{n=k}^{\infty} \lambda^-\Big(A_n\Big) < \sum_{n=k}^{\infty} 1/2^n \;=\; 1/2^{k-1}, \end{aligned}$$

by Equation $**$. But k is arbitrary here and so this tells us that $\lambda^-(A) = 0$.

Also, by Equation $***$,

$$\lambda^+\Big(\bigcap_{n=k}^{\infty} A_n^C\Big) \;\leq\; \lambda^+\Big(A_n^C\Big) \;<\; 1/2^n,$$

for all $n \geq k$*. Finally, since the sets* $\cap_{n \geq k} A_n^C$ *are increasing, we have*

$$\lim_k \lambda^+\Big(\bigcap_{n=k}^{\infty} A_n^C\Big) \;=\; \lambda^+\Big(X \setminus A\Big).$$

We thus conclude $\lambda^+(X \setminus A) = 0$*. Now set* $B = X \setminus A$.
It remains to show that A *is a positive set and* B *is a negative set. Let* E *be measurable. Then* $E \cap A$ *is contained in* A *and so*

$$0 \leq \lambda^-(E \cap A) \;\leq\; \lambda^-(A) \;=\; 0.$$

Then, by the Jordan Decomposition of λ*, we see*

$$\lambda(E \cap A) \;=\; \lambda^+(E \cap A) \;\geq\; 0.$$

This shows that A *is a positive set. A similar argument shows* B *is a negative set.* ■

We can use the Hahn Decomposition to characterize λ^+ and λ^- is a new way.

Lemma 15.1.3. The Hahn Decomposition Characterization of a Charge
Let (A, B) *be a Hahn Decomposition for the charge* λ *on* $(X, \mathcal{S})$*. Then, if* E *is measurable,* $\lambda^+(E) = \lambda(E \cap A)$ *and* $\lambda^-(E) = -\lambda(E \cap B)$*.*

Proof. *Let* D *be a measurable subset of* $E \cap A$*. Then* $\lambda(D) \geq 0$ *by the definition of the positive set* A*. Since* λ *is countably additive, we then have*

$$\begin{aligned}\lambda\Big(E \cap A\Big) &= \lambda\Big((E \cap A) \cap D\Big) + \lambda\Big((E \cap A) \cap D^C\Big)\\ &= \lambda\Big(D\Big) + \lambda\Big((E \cap A) \cap D^C\Big)\end{aligned}$$

But the second set is contained in $E \cap A$ *and so its* λ *measure is non negative. Hence, we can overestimate the left hand side as*

$$\lambda\Big(E \cap A\Big) \;\geq\; \lambda\Big(D\Big) \geq 0.$$

Since this is true for all subsets D*, the definition of* λ^+ *implies* $\lambda^+(E \cap A) \leq \lambda(E \cap A)$*. Now,*

$$\lambda^+(E) \;=\; \lambda^+(E \cap A) + \lambda^+(E \cap B).$$

If F *is a measurable subset of By the definition of* $E \cap B$*, then* $\lambda(F) \leq 0$ *and so* $\sup\{\lambda(F)\} \leq 0$*. This tells us* $\lambda^+(E \cap B) = 0$*. Thus, we have established that* $\lambda^+(E) = \lambda^+(E \cap A)$*. and so* $\lambda^+(E) \leq \lambda(E \cap A)$*.*

The reverse inequality is easier. Since $E \cap A$ is a measurable subset of E, the definition of λ^+ implies $\lambda(E \cap A) \le \lambda^+(E)$. Combining these results, we have $\lambda^+(E) = \lambda(E \cap A)$ as desired.
A similar argument shows that $\lambda^-(E) = -\lambda(E \cap B)$. ■

15.2 The Variation Of A Charge

A charge λ has associated with it a concept that is very similar to that of the variation of a function. We now define the variation of a charge.

Definition 15.2.1. The Variation of a Charge

Let $(X, \mathcal{S})$ be a measure space and λ be a charge on $\mathcal{S}$. For a measurable set E, a mesh in E is a finite collection of disjoint measurable sets inside E, $\{E_1, \ldots, E_n\}$ for some positive integer n. Define the mapping V_λ by

$$V_\lambda(E) \quad = \quad \sup\, \{ \sum_i |\lambda(E_i)| \,|\, \{E_i\} \text{ is a mesh in } E\}$$

where we interpret the sum as being over the finite number of sets in the given mesh. We say $V_\lambda(E)$ is the total variation of λ on E and V_λ is the total variation of λ.

Theorem 15.2.1. The Variation of a Charge is a Measure

Let $(X, \mathcal{S})$ be a measure space and λ be a finite charge on $\mathcal{S}$. Then V_λ is a measure on $\mathcal{S}$.

Proof. *Given a measurable set E, the Jordan Decomposition of λ implies that for a mesh $\{E_1, \ldots, E_n\}$ in E, $|\lambda(E_i)| \le \lambda^+(E_i) + \lambda^-(E_i)$. Hence, since λ^+ and λ^- are measures and countably additive, we have*

$$\begin{aligned} \sum_i |\lambda(E_i)| \quad &\le \quad \sum_i \lambda^+(E_i) \,+\, \sum_i \lambda^-(E_i) \\ &\le \quad \lambda^+(E) \,+\, \lambda^-(E) \,<\, \infty \end{aligned}$$

since λ^+ and λ^- are both finite. We conclude V_λ is a finite mapping.
Since the only mesh in $\emptyset$ is $\emptyset$ itself, we see $V_\lambda(\emptyset) = 0$. It remains to show countable additivity. Let (E_n) be a countable disjoint family in $\mathcal{S}$ and let E be their union. Let $\{A_1, \ldots, A_p\}$ be a mesh in E. Then each A_i is inside E and they are pairwise disjoint. Let $A_{i\,n} = A_i \cap E_n$. Note A_i is the union of the sets $A_{i\,n}$. Then it is easy to see $\{A_{1n}, \ldots, A_{pn}\}$ is a mesh in E_n. For convenience, call this mesh M_n. Then

$$\sum_{i=1}^{p} |\lambda(A_i)| \quad = \quad \sum_{i=1}^{p} \sum_n |\lambda(A_{i\,n}| \,=\, \sum_n \left(\sum_{i=1}^{p} |\lambda(A_{i\,n}| \right).$$

The term in parenthesis is the sum over the mesh M_n of E_n. By definition, this is bounded above by $V_\lambda(E_n)$. Thus, we must have

$$\sum_{i=1}^{p} |\lambda(A_i)| \quad \le \quad \sum_n V_\lambda(A_i).$$

To get the other inequality, we apply the Supremum Tolerance Lemma to the definition of $V_\lambda(E_n)$ to find meshes

$$M_n^\epsilon = \{A_{1\,n}^\epsilon, \ldots, A_{p_n\,n}^\epsilon\},$$

where p_n is a positive integer, so that

$$V_\lambda(E_n) \quad < \quad \sum_{i=1}^{p_n} |\lambda(A_{i\,n}^\epsilon)| \;+\; \epsilon/2^n.$$

It follows that the union of a finite number of these meshes is a mesh of E. For each positive integer N, let

$$\mathcal{M}_N \quad = \quad \bigcup_{i=1}^{N} M_i^\epsilon$$

denote this mesh. Then,

$$\sum_{n=1}^{N} V_\lambda(E_n) \quad < \quad \sum_{n=1}^{N} \left(\sum_{i=1}^{p_n} |\lambda(A_{i\,n}^\epsilon)| \;+\; \epsilon/2^n \right).$$

The first double sum corresponds to summing over a mesh of E and so by definition, we have

$$\sum_{n=1}^{N} V_\lambda(E_n) \quad < \quad V_\lambda(E) \;+\; \sum_{n=1}^{N} \epsilon/2^n \;\leq\; V_\lambda(E) \;+\; \sum_{n=1}^{\infty} \epsilon/2^n \;=\; V_\lambda(E) \;+\; \epsilon.$$

Since N is arbitrary, we see the sequence of partial sums on the left hand side converges to a finite limit. Thus,

$$\sum_{n=1}^{\infty} V_\lambda(E_n) \quad \leq \quad V_\lambda(E) \;+\; \epsilon.$$

Since ϵ is arbitrary, the other desired inequality follows. ■

Theorem 15.2.2. $V_\lambda \;=\; \lambda^+ \;+\; \lambda^-$

Let $(X, \mathcal{S})$ be a measure space and λ be a finite charge on $\mathcal{S}$. Then $V_\lambda \;=\; \lambda^+ \;+\; \lambda^-$.

Proof. *Choose a measurable set E and let $\epsilon > 0$ be chosen. Then, by the Supremum Tolerance Lemma, there is a mesh $M^\epsilon = \{A_1^\epsilon, \ldots, A_p^\epsilon$ so that*

$$V_\lambda(E) \;-\; \epsilon \quad < \quad \sum_i |\lambda(A_i^\epsilon)| \;\leq\; V_\lambda(E).$$

Let F be the set of indices i in the mesh above where $\lambda(A_i^\epsilon) \;\geq\; 0$ and G be the other indices where $\lambda(A_i^\epsilon) < 0$. Let $\mathcal{F}$ be the union over the indices in F and $\mathcal{G}$ be the union over the indices in G. Note we *have*

$$\begin{aligned} V_\lambda(E) \, - \, \epsilon &< \sum_i |\lambda(A_i^\epsilon)| \\ &= \sum_F |\lambda(A_i^\epsilon)| \, + \, \sum_F |\lambda(A_i^\epsilon)|. \end{aligned}$$

Now in F,

$$\begin{aligned} |\lambda(A_i^\epsilon)| &= \lambda^+(A_i^\epsilon) - \lambda^-(A_i^\epsilon) \\ &\leq \lambda^+(A_i^\epsilon), \end{aligned}$$

and in G,

$$\begin{aligned} |\lambda(A_i^\epsilon)| &= \lambda^-(A_i^\epsilon) - \lambda^+(A_i^\epsilon) \\ &\leq \lambda^-(A_i^\epsilon). \end{aligned}$$

Thus, we can say

$$\begin{aligned} V_\lambda(E) \, - \, \epsilon &\leq \sum_F \lambda^+(A_i^\epsilon) \, + \, \sum_G \lambda^-(A_i^\epsilon) \\ &= \lambda^+(\mathcal{F}) \, + \, \lambda^-(\mathcal{G}) \\ &\leq \lambda^+(\mathcal{E}) \, + \, \lambda^-(\mathcal{E}). \end{aligned}$$

Thus, for all $\epsilon > 0$, we have

$$V_\lambda(E) \quad \leq \quad \lambda^+(\mathcal{E}) \, + \, \lambda^-(\mathcal{E}) \, + \, \epsilon.$$

This implies

$$V_\lambda(E) \quad \leq \quad \lambda^+(\mathcal{E}) \, + \, \lambda^-(\mathcal{E}).$$

To prove the reverse, note if $A \subseteq E$ for $E \in \mathcal{S}$, then A itself is a mesh (a pretty simple one, of course) and so $|\lambda(A)| \leq V_\lambda(A)$. Further, $\lambda(E) = \lambda(A) + \lambda(E \setminus A)$. Thus, we have

$$\begin{aligned} 2\,\lambda(A) &\leq \lambda(A) \, + \, |\lambda(A)| \, \leq \, \lambda(E) \, - \, \lambda(E \setminus A) \, + \, |\lambda(A)| \\ &\leq \lambda(E) \, + \, |\lambda(E \setminus A)| \, + \, |\lambda(A)| \end{aligned}$$

But the collection $\{A, E \setminus A\}$ is a mesh for E and so

$$2\,\lambda(A) \quad \leq \quad \lambda(E) \, + \, V_\lambda(E).$$

Next, using the definition of λ^+, we find

$$2\,\lambda^+(E) \quad \leq \quad \lambda(E) \, + \, V_\lambda(E).$$

Finally, using the Jordan Decomposition of λ, we obtain

$$2\,\lambda^+(E) \;\; \leq \;\; \lambda^+(E) \,-\, \lambda^-(E) \,+\, V_\lambda(E).$$

This immediately leads to $\lambda^+(E) - \lambda^-(E) \leq V_\lambda(E)$. ∎

15.3 Absolute Continuity Of Charges

Now we are ready to look at absolute continuity in the context of charges.

Definition 15.3.1. Absolute Continuity Of Charges

Let $(X, \mathcal{S}, \mu)$ be a measurable space and let λ be a charge on $\mathcal{S}$. Then λ is said to be absolutely continuous with respect to μ if whenever E is a measurable set with $\mu(E) = 0$, then $\lambda(E) = 0$ also. We write this as $\lambda \ll \mu$. The set of all charges that are absolutely continuous with respect to μ is denoted by $AC[\mu]$.

There is an intimate relationship between the absolute continuity of V_λ, λ, λ^+ and λ^-; essentially, one implies all the others.

Theorem 15.3.1. Equivalent Absolute Continuity Conditions For Charges

Let $(X, \mathcal{S}, \mu)$ be a measurable space. Then for the statements (1): λ^+ and λ^- are in $AC[\mu]$,
(2): V_λ is in $AC[\mu]$, and
(3): λ is in $AC[\mu]$, we have (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3).

Proof.
(1) $\rightarrow$ (2): if $\mu(E) = 0$, then $\lambda^+(E)$ and $\lambda^-(E)$ are also zero by assumption. Applying the Jordan Decomposition of λ, we see $\lambda(E) = 0$ too. Hence, λ is in $AC[\mu]$.
(2) $\rightarrow$ (3): if $\mu(E) = 0$, then $V_\lambda(E) = 0$. But, by Theorem 15.2.2, we have both $\lambda^+(E)$ and $\lambda^-(E)$ are zero. Then, applying the Jordan Decomposition again, we have $\lambda(E) = 0$. This tells us λ is absolutely continuous with respect to μ.
(3) $\rightarrow$ (1): Let (A, B) be a Hahn Decomposition of X due to λ. If $\mu(E) = 0$, then $\lambda(E) = 0$ by assumption. Thus, $\lambda(E \cap A) = \lambda(E \cap B) = 0$ as well. By Lemma 15.1.3, we then have that $\lambda^+(E) = \lambda^-(E) = 0$ showing that (1) holds. ∎

There is another characterization of absolute continuity that is useful.

Lemma 15.3.2. $\epsilon - \delta$ Version Of Absolute Continuity Of a Charge

Let λ be a finite charge of $\mathcal{S}$. Then

$$\lambda \ll \mu \quad \Leftrightarrow \quad \forall\,\epsilon \,>\, 0, \exists\,\delta \,>\, 0 \;\ni\; |\lambda(E)| \,<\, \epsilon \text{ for measurable } E \text{ with } \mu(E) \,<\, \delta.$$

Proof.
($\Rightarrow$): If λ is absolutely continuous with respect to μ, then by Theorem 15.3.1 (the previous result) V_λ is also in AC[μ]. We will prove this by contradiction. Assume the desired implication does not hold for V_λ. Then, there is a positive ϵ so that for all n, there is a measurable set E_n with $\mu(E_n) < 1/2^n$ and $V_\lambda(E_n) \geq \epsilon$.

Let

$$\begin{aligned} G_n &= \bigcup_{k=n}^{\infty} E_k, \\ G &= \bigcap_n G_n. \end{aligned}$$

Then,

$$\mu(G) \leq \mu(G_n) \leq \sum_{k=n}^{\infty} E_k < \sum_{k=n}^{\infty} 1/2^k = 1/2^{n-1}.$$

Since this holds for all n, this implies $\mu(G) = 0$. Since V_λ is in AC[μ], we then have $V_\lambda(G) = 0$. But

$$V_\lambda(G) = \lim_n V_\lambda(G_n) \geq \epsilon.$$

This contradiction implies that our assumption that the right hand side did not hold must be false. Hence, the condition holds for V_λ. It is easy to see that since $V_\lambda = \lambda^+ + \lambda^-$, that the condition holds for them also. This then implies the condition holds for $\lambda = \lambda^+ - \lambda^-$.
($\Leftarrow$): We assume the condition on the right hand side holds. Now let (A, B) be a Hahn Decomposition for X with respect to λ. In particular, if $\mu(E) = 0$, then $\mu(E \cap A) = 0$ also. The condition then implies $\lambda(E \cap A) < \epsilon$. However, the choice of ϵ is arbitrary which then implies $|\lambda(E \cap A)| = 0$. But the absolute values are unnecessary as λ is non negative on A. We conclude $\lambda^+(E) = \lambda(E \cap A) = 0$. A similar argument then shows $\lambda^-(E) = -\lambda(E \cap B) = 0$. This tells us $\lambda(E) = 0$ by the Jordan Decomposition.
■

Lemma 15.3.3. The Absolute Continuity Of The Integral
Let $(X, \mathcal{S}, \mu)$ be a measure space and f be a summable function. Define the map λ by $\lambda(E) = \int_E f \, d\mu$ for all measurable E. Then, λ is a charge with

$$\lambda^+(E) = \int_E f^+ \, d\mu, \quad \lambda^-(E) = -\int_E f^- \, d\mu.$$

Moreover, if $P_f = \{x \mid f(x) \geq 0\}$ and $N_f = P_f^C$, then (P_f, N_f) is a Hahn Decomposition for X with respect to λ. Finally, since $\lambda \ll \mu$, we know for all positive ϵ, there is a positive δ, so that if E is a measurable set with $\mu(E) < \delta$, then

$$\left| \int_E f \, d\mu \right| < \epsilon,$$

Proof. *It is easy to see that* $\nu_1 = \int_E f^+ \, d\mu$ *and* $\nu_2 = \int_E f^- \, d\mu$ *define measures and that* $\lambda = \nu_1 - \nu_2$*. Hence,* λ *is a charge which is absolutely continuous with respect to* μ*. It is also easy to see that* (P_f, N_f) *is a Hahn Decomposition for* λ*. Now if* B *is measurable and contained in the measurable set* E*, we have*

$$\begin{aligned} \lambda(B) &= \int_{B \cap P_f} f^+ \, d\mu - \int_{B \cap N_f} f^+ \, d\mu \\ &\leq \int_{B \cap P_f} f^+ \, d\mu \\ &\leq \int_{E \cap P_f} f^+ \, d\mu. \end{aligned}$$

Next, note that $\int_{E \cap P_f} f^+ \, d\mu = \int_E f^+ \, d\mu$ *because the portion of* E *that lies in* N_f *does not contribute to the value of the integral. Thus, for any* $B \subseteq E$*, we have*

$$\lambda(B) \leq \int_E f^+ \, d\mu = \nu_1(E).$$

The definition of λ^+ *then implies two things: first, the inequality above tells us* $\lambda^+(E) \leq \nu_1(E)$ *and second, since* $E \cap P_f$ *is a subset of* E*, we know* $\lambda(E \cap P_f) \leq \lambda^+(E)$*. However,* $\lambda(E \cap P_f) = \nu_1(E)$ *and hence,* $\nu_1(E) \leq \lambda^+(E)$ *also. Combining, we have* $\lambda^+(E) = \nu_1(E)$*.*

A similar argument shows that $\lambda^-(E) = \nu_2(E)$*.*

The last statement of the proposition follows immediately from Lemma 15.3.2. ■

15.4 The Radon - Nikodym Theorem

From our work above, culminating in Lemma 15.3.3, we know that integrals of summable functions define charges which are absolutely continuous with respect to the measure we are using for the integration. The converse of this is that if a measure is absolutely continuous, we can find a summable function so that the measure can be found by integration. That is if $\lambda \ll \mu$, there exists f summable so that $\lambda(E) = \int f \, d\mu$. This result is called the Radon - Nikodym theorem and as you might expect, its proof requires some complicated technicalities to be addressed. Hence, we begin with a lemma.

Lemma 15.4.1. Radon - Nikodym Technical Lemma

Let $(X, \mathcal{S}, \mu)$ *be a measurable space with* $\mu(X)$ *finite. Let* λ *be a measure which is finite with* $\lambda(X) > 0$ *and* $\lambda \ll \mu$*. Then there is a positive* ϵ *and a measurable set* A *with* $\mu(A) > 0$ *so that*

$$\epsilon \, \mu(E \cap A) \leq \lambda(E \cap A), \forall \, E \in \mathcal{S}.$$

Proof. *Pick a fixed* $\epsilon > 0$ *and assume the set* A *exists. Let* $\nu = \lambda - \epsilon \, \mu$*. Then,* ν *is a finite charge also. Note, our assumption tells us that*

$$\nu(B) = \lambda(B) - \epsilon \, \mu(B) \geq 0,$$

for all measurable subsets B of A. Hence, by the definition of ν^-, we must have that $-\nu^-(A) \geq 0$ or $\nu^-(A) \leq 0$. But ν^- is always non negative. Combining, we have $\nu^-(A) = 0$. This gives us some clues as to how we can find the desired A. Note if (A, B) is a Hahn Decomposition for ν, then we have this desired inequality, $\nu^-(A) = 0$. So, we need to find a positive value of ϵ^* so that when (A, B) is a Hahn Decomposition of

$$\nu^*(A) \quad = \quad \lambda(A) \, - \, \epsilon^* \, \mu(A),$$

we find $\nu^*(A) > 0$.

To do this, for $\epsilon = 1/n$, let (A_n, B_n) be a Hahn Decomposition for $\nu_n = \lambda - (1/n)\,\mu$. Let $G = \cup_n A_n$ and $H = \cap_n B_n$. We also know $A_n \cup B_n = X$ and $A_n \cap B_n = \emptyset$ for all n. Further,

$$H^C \quad = \quad \left(\bigcap_n B_n \right)^C = \bigcup_n B_n^C = \bigcup_n A_n \; = \; G.$$

We conclude $X = G \cup H$; it is easy to see $G \cap H = \emptyset$. Now, $H \subseteq B_n$ for all n, so $\nu_n(H) = -\nu_n^-(H) \leq 0$ as B_n is a negative set. Hence, we can say

$$\lambda(H) \, - \, (1/n)\,\mu(H) \quad \leq \quad 0$$

which implies $\lambda(H) \, \leq \, (1/n)\,\mu(H)$ for all n. Since λ is a measure, we then have

$$0 \; \leq \; \lambda(H) \; \leq \; \mu(H)/n$$

which implies by the arbitrariness of n that $\lambda(H) = 0$. Hence,

$$\lambda(X) \quad = \quad \lambda(G) \, + \, \lambda(H) \; = \; \lambda(G).$$

Thus, $\lambda(G) > 0$ as $\lambda(X) > 0$. Since $\lambda \ll \mu$, it then follows that $\mu(G) > 0$ also. Since $G = \cup_n A_n$, it must be true that there is at least one n with $\mu(A_n) > 0$. Call this index N. Then, $\nu_N(E \cap A_N) \geq 0$ as A_N is a positive set for ν_N. This implies

$$\lambda(E \, \cap \, A_N) \, - \, \frac{\mu(E \, \cap \, A_N)}{N} \quad \geq \quad 0,$$

which is the result we seek using $A = A_N$ and $\epsilon = 1/N$. ∎

Theorem 15.4.2. The Radon - Nikodym Theorem

Let $(X, \mathcal{S}, \mu)$ be a measurable space with μ σ - finite. Let λ be a charge with $\lambda \ll \mu$. Then, there is a summable function f so that

$$\lambda(E) \quad = \quad \int_E f \, d\mu$$

for all measurable E. Moreover, if g is another summable function which satisfies this equality, then $f = g$ μ a.e. The summable function f is called the Radon - Nikodym derivative of λ with respect to μ and is often denoted by the usual derivative symbol: $f = \frac{d\lambda}{d\mu}$. Hence, this equality is often written

$$\lambda(E) \quad = \quad \int_E \frac{d\lambda}{d\mu} \, d\mu$$

Proof. *We will do this in three steps.*
Step 1: *We assume $\mu(X)$ is finite and λ is a finite measure.*
Step 2: *We assume μ is σ - finite and λ is a finite measure.*
Step 3: *We assume μ is σ - finite and λ is a finite charge.*

As is usual, the proof of **Step 1** *is the hardest.*
Proof Step 1: *Let*

$$\mathcal{F} \quad = \quad \{ \, f : X \to \Re \, | \, f \, \geq \, 0, \, f \;\; \textit{summable and} \; \int_E f \, d\mu \, \leq \, \lambda(E), \, \forall \, E \, \in \, \mathcal{S}\}.$$

Note since $f = 0_X$, $\mathcal{F}$ is nonempty. From the definition of $\mathcal{F}$, we see $\int_X f \, d\mu \leq \lambda(X) < \infty$ for all f in $\mathcal{F}$. Hence,

$$c \quad = \quad \sup_{f \in \mathcal{F}} \int_X f \, d\mu \, < \, \infty.$$

We will find a particular $f \in \mathcal{F}$ so that $c = \int_X f \, d\mu$. Let $(f_n) \subseteq \mathcal{F}$ be a minimizing sequence: i.e. $\int_X f_n \, d\mu \to c$. We will assume without loss of generality that each f_n is finite everywhere as the set of points where all are infinite is a set of measure zero. Now, there are details that should be addressed in that statement, but we have gone through those sort of manipulations many times before. As an exercise, you should go through them again on scratch paper for yourself. With that said, we will define a new sequence of finite functions (g_n) by

$$\begin{aligned} g_n \quad &= \quad f_1 \, \vee \, f_2 \, \vee \, \ldots \, \vee \, f_n \\ &= \quad \max \{f_1, \ldots, f_n\}. \end{aligned}$$

This is a pointwise operation and it is clear that (g_n) is an increasing sequence of non negative functions. Since f_1 and f_2 are summable, let A be the set of points where $f_1 > f_2$. Then,

$$\begin{aligned} \int_X f_1 \vee f_2 \, d\mu &= \int_A f_1 \, d\mu + \int_{A^C} f_2 \, d\mu \\ &\leq \int_X f_1 \, d\mu + \int_X f_2 \, d\mu. \end{aligned}$$

This tells us $f_1 \, f_2$ is summable also. A simple induction argument then tells us g_n is summable for all n.

Is $g_n \in \mathcal{F}$? Let E be measurable. Define the measurable sets (E_n) by

$$\begin{aligned} E_1 &= \{ x \,|\, g_n(x) = f_1(x)\} \cap E, \\ E_2 &= \{ x \,|\, g_n(x) = f_2(x)\} \cap (E \setminus E_1), \\ &\vdots \\ E_n &= \{ x \,|\, g_n(x) = f_n(x)\} \cap (E \setminus \cup_{i=1}^{n-1} E_i). \end{aligned}$$

Then, it is clear $E = \cup_i E_i$, each E_i is disjoint from the others and $g_n(x) = f_i(x)$ on E_i. Thus, since each f_i is in $\mathcal{F}$, we have

$$\begin{aligned} \int_E g_n \, d\mu &= \sum_{i=1}^n \int_{E_i} f_i \, d\mu \\ &\leq \sum_{i=1}^n \lambda(E_i) \;=\; \lambda(\cup_{i=1}^n E_i) \\ &= \lambda(E). \end{aligned}$$

We conclude each g_n is in $\mathcal{F}$ for all n. Next, if $g = \sup g_n$, then $g_n \uparrow g$ and

$$\int_E g_n \, d\mu \leq \lambda(E) \quad \leq \quad \lambda(X)$$

for all n. Now apply the Monotone Convergence Theorem to see g is summable and

$$\int_E g_n \, d\mu \to \int_E g \, d\mu \;\leq\; \lambda(E).$$

Let's define f by

$$f(x) \quad = \quad \begin{cases} g(x) & g(x) \;<\; \infty \\ ,0 & g(x) \;=\; \infty. \end{cases}$$

Since g is summable, the set of points where it takes on the value ∞ is a set of measure 0. Thus, $f = g$ μ a.e. and f is measurable. It is easy to see f is in $\mathcal{F}$.

Moreover, since $f_n \leq g_n$, we have

$$\begin{aligned} c &= \lim_n \int_X f_n \, d\mu \\ &\leq \lim_n \int_X g_n \, d\mu \leq c \end{aligned}$$

because $g_n \in \mathcal{F}$. Thus,

$$c = \lim_n \int_X g_n \, d\mu = \int_X g \, d\mu.$$

This immediately tells us that $\int_X f \, d\mu = c$ with $f \in \mathcal{F}$.

Next, define $m : \mathcal{S} \to \Re$ by

$$m(E) = \lambda(E) - \int_E f \, d\mu,$$

for all measurable E. It is straightforward to show m is difference of two measures and hence is a finite charge. Also, since f is in $\mathcal{F}$, we see m is non negative and thus is a measure. In addition, since $\lambda \ll \mu$ and the measure defined by $\int_E f \, d\mu$ is also absolutely continuous with respect to μ, we have that $m \ll \mu$ too. Now if $m(X) = 0$, this would imply, since $m(E) \leq m(X)$, that

$$0 \leq \lambda(E) - \int_E f \, d\mu \leq m(X) = 0.$$

But this says $\lambda(E) = \int_E f \, d\mu$ for all measurable E which is the result we seek.

Hence, it suffices to show $m(X) = 0$. We will do this by contradiction. Assume $m(X) > 0$. Now apply Lemma 15.4.1 to conclude there is a positive ϵ and measurable set A so that $\mu(A) > 0$ and

$$\epsilon \, \mu(E \cap A) \leq m(E \cap A), \qquad (*)$$

for all measurable E. Define a new function h using Equation $*$ by $h = f + \epsilon \, I_A$. Then for a given measurable E, we have

$$\begin{aligned} \int_E h \, d\mu &= \int_E f \, d\mu + \epsilon \, \mu(E \cap A) \\ &\leq \int_E f \, d\mu + m(E \cap A) \end{aligned}$$

by Equation $*$. Now replace m by its definition to find

$$\begin{aligned} \int_E h \, d\mu &\leq \int_E f \, d\mu + \lambda(E \cap A) - \int_{E \cap A} f \, d\mu \\ &= \int_{E \cap A^C} f \, d\mu + \lambda(E \cap A). \end{aligned}$$

Finally, use the fact that f is in $\mathcal{F}$ to conclude

$$\int_E h\, d\mu \;\; \leq \;\; \lambda(E \cap A^C) \, + \, \lambda(E \cap A) \; = \; \lambda(E).$$

This shows that h is in $\mathcal{F}$. However,

$$\int_X h\, d\mu \;\; = \;\; \int_X f\, d\mu \, + \, \epsilon\, \mu(A) \; > \; c!$$

which is our contradiction. This completes the proof of **Step 1**.

Proof Step 2: *Now μ is σ finite. This means there is a countable sequence of disjoint measurable sets (X_n) with $\mu(X_n)$ finite for each n and we can write $X = \cup_n X_n$. Let $\mathcal{S}_n$ be the σ - algebra of subsets of X_n given by $\mathcal{S} \cap X_n$. By* **Step 1**, *there are summable non negative functions f_n so that*

$$\lambda(F) \;\; = \;\; \int_F f_n\, d\, \mu,$$

for each F in $\mathcal{S}_n$. Now define f by $f(x) = f_n(x)$ when $x \in X_n$. This is a well - defined function and it is easy to see f is measurable. If E is measurable, then $E = \cup_n E \cap X_n$, $E = \cup_n E \cap X_n$ and

$$\int_E f\, d\, \mu \;\; = \;\; \int_{\cup_n \, E \cap X_n} f\, d\, \mu.$$

Then, for any n,

$$\begin{aligned} \int_{\cup_{i=1}^n \, E \cap X_i} f\, d\, \mu \;\; &= \;\; \sum_{i=1}^n \int_{E \cap X_i} f\, d\, \mu = \sum_{i=1}^n \int_{E \cap X_i} f_n\, d\, \mu \\ &= \;\; \sum_{i=1}^n \lambda(E \cap X_i) \; = \; \lambda(\cup_{i=1}^n \, E \cap X_i) \; \leq \; \lambda(E), \end{aligned}$$

which is a finite number. Hence, the series of non negative terms $\sum_n \int_{E \cap X_n} f\, d\, \mu$ converges and

$$\int_E f\, d\, \mu \;\; = \;\; \sum_n \int_{\cup_n \, E \cap X_n} f_n\, d\, \mu \; = \; \lambda(\cup_n \, E \cap X_n) \; = \; \lambda(E).$$

This establishes the result for **Step 2**.

Proof Step 3: *Here, we have μ is σ - finite and λ is a finite charge. By the Jordan Decomposition of λ, we can write*

$$\lambda(E) \;\; = \;\; \lambda^+(E) \, - \, \lambda^-(E),$$

for all measurable E. Now apply **Step 2** *to find non negative summable functions f^+ and f^- so that*

$$\begin{aligned} \lambda^+(E) \;\; &= \;\; \int_E f^+\, d\, \mu, \\ \lambda^-(E) \;\; &= \;\; \int_E f^-\, d\, \mu. \end{aligned}$$

Let $f = f^+ - f^-$ and we are done with the proof of **Step 3***.*

Finally, it is clear from the proof above, that the Radon - Nikodym derivative of λ with respect to μ, is unique up to redefinition on a set of μ measure 0. ∎

15.5 The Lebesgue Decomposition of a Measure

Definition 15.5.1. Singular Measures

Let $(X, \mathcal{S}, \mu)$ be a measure space and let λ be a charge on $\mathcal{S}$. Assume there is a decomposition of X into disjoint measurable subsets U and V ($X = U \cup V$ and $U \cap V = \emptyset$) so that $\mu(U) = 0$ and $\lambda(E \cap V) = 0$ for all measurable subsets E of V. In this case, we say λ is perpendicular to μ and write $\lambda \perp \mu$.

Comment 15.5.1. *If $\lambda \perp \mu$, let (U, V) be a decomposition of X associated with the singular measure λ. We then know that $\mu(U) = 0$ and $\lambda(E \cap V) = 0$ for all measurable E. Note, if E is measurable, then*

$$E = \Big(E \cap U\Big) \cup \Big(E \cap V\Big).$$

Thus,

$$\lambda(E) = \lambda\Big(E \cap U\Big) + \lambda\Big(E \cap V\Big) = \lambda\Big(E \cap U\Big).$$

Further,

$$\mu(E) = \mu\Big(E \cap U\Big) + \mu\Big(E \cap V\Big) = \mu\Big(E \cap V\Big).$$

Comment 15.5.2. *If $\lambda \perp \mu$ with $\lambda \neq 0$, then there is a measurable set E so that $\lambda(E \cap U) \neq 0$. But for this same set $\mu(E \cap U) = 0$ as $E \cap A$ is a subset of U. Thus, $\lambda \not\ll \mu$.*

Comment 15.5.3. *If $\lambda \perp \mu$ and $\lambda \ll \mu$, then for any measurable set E, we have $\lambda(E) = \lambda(E \cap U)$. But, since $\mu(E \cap U) = 0$, we must have $\lambda(E \cap U) = 0$ because $\lambda \ll \mu$. Thus, $\lambda = 0$.*

Comment 15.5.4. *It is easy to prove that $\lambda \perp \mu$ implies $V_\lambda \perp \mu$, $\lambda^+ \perp \mu$ and $\lambda^- \perp \mu$. Also, if $\lambda^+ \perp \mu$ and $\lambda^- \perp \mu$, this implies $\lambda \perp \mu$.*

Theorem 15.5.1. Lebesgue Decomposition Theorem

Let $(X, \mathcal{S}, \mu)$ be a σ - finite measure space. Let λ be a finite charge on $\mathcal{S}$. Then, there are two unique finite measures, $\lambda_{ac} \ll \mu$ and $\lambda_p \perp \mu$ such that $\lambda = \lambda_{ac} + \lambda_p$.

Proof. *We will prove this result in four steps.*
Step 1*: λ and μ are finite measures.*

Step 2: *μ is a σ - finite measure and λ is a finite measure.*

Step 3: *μ is a σ - finite measure and λ is a finite charge.*

Step 4: *The decomposition is unique.*

Proof Step 1: *As is usual, this is the most difficult step. We can see, in this case, that $\lambda + \mu$ is a measure. Note that $(\lambda + \mu)(E) = 0$ implies that $\lambda(E)$ is 0 too. Hence, $\lambda \ll (\lambda + \mu)$. By the Radon - Nikodym Theorem, there is then a non negative $\lambda + \mu$ summable f so that for any measurable E,*

$$\lambda(E) \quad = \quad \int_E f \, d\,(\lambda \, + \, \mu).$$

Hence, f is μ and λ summable as well and

$$\lambda(E) \quad = \quad \int_E f \, d\,\lambda \, + \, \int_E f \, d\,\mu).$$

Let

$$\begin{aligned} A_1 &= \{x \mid f(x) \, = \, 1\}, \\ A_2 &= \{x \mid f(x) \, > \, 1\}, \quad \textit{and} \\ B &= \{x \mid f(x) \, < \, 1\}. \end{aligned}$$

Also, for each n, let

$$E_n \quad = \quad \{x \mid f(x) \, \geq \, 1 \, + \, 1/n\}.$$

Then, we see immediately $A_2 = \cup_n E_n$ and $X = A \cup B$. Now, we also have

$$\begin{aligned} \lambda(E_n) &= \int_{E_n} f \, d\,(\lambda \, + \, \mu) \\ &\geq (1 + 1/n) \left(\lambda(E_n) \, + \, \mu(E_n) \right). \end{aligned}$$

This implies $\lambda(E_n) \, \geq \, (1 + 1/n) \, \lambda(E_n)$ which tells us $\lambda(E_n) \, \leq \, 0$. But since λ is a measure, this forces $\lambda(E_n) = 0$. From the same inequality, we also have $\lambda(E_n) \geq \lambda(E_n) \, + \, \mu(E_n)$. which forces $\mu(E_n) = 0$ too.
Next, note the sequence of sets (E_n) increases to A_2 and so

$$\begin{aligned} \lim_n \, \mu(E_n) &= \mu(A_2), \\ \lim_n \, \lambda(E_n) &= \lambda(A_2). \end{aligned}$$

Since $\mu(E_n) = \lambda(E_n) = 0$ for all n, we conclude $\mu(A_2) = \lambda(A_2) = 0$.

Also,

$$\begin{aligned}
\lambda(A_1) &= \int_{A_1} f \, d\,(\lambda + \mu) \\
&= \int_{A_1} 1 \, d\,(\lambda + \mu) \\
&= \mu(A_1) + \lambda(A_1),
\end{aligned}$$

which implies $\mu(A_1) = 0$. Let $A = A_1 \cup A_2$. Then, the above remarks imply $\mu(A) = 0$. We now suspect that A and B will gives us the decomposition of X which will allow us to construct the measures $\lambda_{ac} \ll \mu$ and $\lambda_p \perp \mu$. Define λ_{ac} and λ_p by

$$\begin{aligned}
\lambda_{ac} &= \lambda(E \cap B), \\
\lambda_p &= \lambda(E \cap A).
\end{aligned}$$

Then,

$$\begin{aligned}
\lambda(E) &= \lambda(E \cap A) + \lambda(E \cap B) \\
&= \lambda_{ac}(E) + \lambda_p(E),
\end{aligned}$$

showing us the we have found a decomposition of λ into two measures.
Is $\lambda_{ac} \ll \mu$? Let $\mu(E) = 0$. Then $\mu(E \cap B) = 0$ as well. Now, we know

$$\begin{aligned}
\lambda(E \cap B) &= \int_{E\cap B} f \, d\,(\lambda + \mu) \\
&= \int_{E\cap B} f \, d\,\lambda + \int_{E\cap B} f \, d\,\mu.
\end{aligned}$$

However, the second integral must be zero since $\mu(E \cap B) = 0$. Thus, we have

$$\lambda(E \cap B) = \int_{E\cap B} f \, d\,\lambda.$$

We also have $\lambda(E \cap B) = \int_{E\cap B} 1 \, d\lambda$ and so

$$\int_{E\cap B} 1 \, d\,\lambda = \int_{E\cap B} f \, d\,\lambda.$$

Thus,

$$\int_{E\cap B} \Big(1 - f\Big) \, d\,\lambda = 0.$$

But on $E\cap B$, $1-f > 0$; hence, we must have $\lambda(E\cap B) = 0$. This means $\lambda_{ac}(E\cap B) = 0$ implying $\lambda_{ac} \ll \mu$.

Is $\lambda_p \perp \mu$? Note, for any measurable E, we have

$$\lambda_p(E \cap B) = \lambda\Big((E \cap B) \cap A\Big) = \lambda(\emptyset) = 0.$$

Thus, $\lambda_p \perp \mu$. In fact, we have shown

$$\lambda(E) \;=\; \int_{E\cap B} f\, d\,\lambda \,+\, \lambda_p(E).$$

Proof Step 2*: Note that once we find a decomposition $X = A \cup B$ with A and B measurable and disjoint satisfying $\mu(A) = 0$ and $\lambda(E \cap B) = 0$ if $\mu(E) = 0$, then we can use the technique in the proof of* **Step 1***. We let $\lambda_{ac}(E) = \lambda(E \cap B)$ and $\lambda_p(E) = \lambda(E \cap A)$. This furnishes the decomposition we seek. Hence, we must find a suitable A and B.*

The measure μ is now σ - finite. Hence, there is a sequence of disjoint measurable sets X_n with $\mu(X_n) < \infty$ and $X = \cup_n X_n$. Let $\mathcal{S}_n$ denote the σ - algebra of subsets $\mathcal{S} \cap X_n$. By **Step 1***, there is a decomposition $X_n = A_n \cup B_n$ of disjoint and measurable sets so that $\mu(A_n) = 0$ and $\lambda(E \cap B_n) = 0$ if $\mu(E) = 0$. Since the sets X_n are mutually disjoint, we know the sequences (A_n) and (A_n) are disjoint also. Let $A = \cup_n A_n$ and $B = A^C$ and note $A^C = \cap_n B_n$. Then, since μ is a measure, we have*

$$\mu(\cup_{i=1}^{n} A_i) \;=\; \sum_{i=1}^{n} \mu(A_i) \;=\; 0$$

for all n. Hence,

$$\mu(A) \;=\; \lim_{n} \mu(\cup_{i=1}^{n} A_i) \;=\; 0.$$

Next, if $mu(E) = 0$, then $\mu(E \cap B_n) = 0$ for all n by the properties of the decomposition (A_n, B_n) of X_n. Since

$$E \cap B \;=\; \cap_n \Big(E \cap B_n\Big),$$

and $\lambda(E \cap B_1)$ is finite, we have

$$\lambda(E \cap B) \;=\; \lim_{n} \lambda(E \cap B_n).$$

However, each $\lambda(E \cap B_n)$ is zero because $\mu(E) = 0$ by assumption. Thus, we conclude $\lambda(E \cap B) = 0$. We then have the A and B we need to construct the decomposition.

Proof Step 3*: The mapping λ is now a finite charge. Let $\lambda = \lambda^+ - \lambda^-$ be the Jordan Decomposition of the charge λ. Applying* **Step 2***, we see there are pairs of measurable sets (A_1, B_1) and (A_2, B_2) so that*

$$X \;=\; A_1 \cup B_1,\; A_1 \cap B_1 \;=\; \emptyset,\; \mu(A_1) \;=\; 0, \mu(E) \;=\; 0 \Rightarrow \lambda^+(E \cap B_1) \;=\; 0,$$

and

$$X \;=\; A_2 \cup B_2,\; A_2 \cap B_2 \;=\; \emptyset, \mu(A_2) \;=\; 0, \mu(E) \;=\; 0 \Rightarrow \lambda^-(E \cap B_2) \;=\; 0.$$

Let $A = A_1 \cup A_2$ and $B = B_1 \cap B_2$. Note $B^C = A$. It is clear then that $\mu(A) = 0$. Finally, if $\mu(E) = 0$, then $\lambda^+(E \cap B_1) = 0$ and $\lambda^-(E \cap B_2) = 0$. This tells us

$$\begin{aligned}
\lambda(E \cap B) &= \lambda^+(E \cap B) - \lambda^-(E \cap B) \\
&= \lambda^+(E \cap B_1 \cap B_2) - \lambda^-(E \cap B_1 \cap B_2) \\
&= \lambda^+\Big((E \cap B_1) \cap B_2\Big) - \lambda^-\Big((E \cap B_2) \cap B_1\Big).
\end{aligned}$$

Both of the terms on the right hand side are then zero because we are computing measures of subsets of a set of measure 0. We conclude $\lambda(E \cap B) = 0$. The decomposition is then

$$\begin{aligned}
\lambda_{ac}(E) &= \lambda(E \cap B) = \Big(\lambda^+ - \lambda^-\Big)(E \cap B), \\
\lambda_p(E) &= \lambda(E \cap A) = \Big(\lambda^+ - \lambda^-\Big)(E \cap A).
\end{aligned}$$

Proof Step 4: *To see this decomposition is unique, assume $\lambda = \lambda_1 + \lambda_2$ and $\lambda_{ac} + \lambda_p$ are two Lebesgue decompositions of λ. Then, $\lambda_{ac} - \lambda_1 = \lambda_2 - \lambda_p$. But since λ_1 and λ_{ac} are both absolutely continuous with respect to μ, it follows that $\lambda_{ac} - \lambda_1 \ll \mu$ also. Further, since both λ_2 and λ_p are singular with respect to μ, we see $\lambda_2 - \lambda_p \perp \mu$. However, $\lambda_{ac} - \lambda_1 = \lambda_2 - \lambda_p$ by assumption and so $\lambda_{ac} - \lambda_1 \ll \mu$ and $\lambda_{ac} - \lambda_1 \perp \mu$. By Comment 15.5.3, this tells us $\lambda_{ac} = \lambda_1$. This then implies $\lambda_2 = \lambda_p$.*

■

15.6 Homework

Exercise 15.6.1. *Let $(X, \mathcal{S})$ be a measurable space and λ is a charge on $\mathcal{S}$. Prove if P_1 and P_2 are positive sets for λ, then $P_1 \cup P_2$ is also a positive set for λ.*

Exercise 15.6.2. *Let $g_1(x) = 2x$, $g_2(x) = I_{[0,\infty)}$, $g_3(x) = x\, I_{[0,\infty)}$ and $g_4(x) = \arctan(x)$. All of these functions generate Borel - Stieljes measures on $\Re$.*

- *(i): Determine which are absolutely continuous with respect to Borel measure. Then, if absolutely continuous with respect to Borel measure, find their Radon - Nikodym derivative.*
- *(ii): Which of these measures are singular with respect to Borel measure?*

Exercise 15.6.3. *Let λ and μ be σ - finite measures on $\mathcal{S}$, a σ - algebra of subsets of a set X. Assume λ is absolutely continuous with respect to μ. If $g \in M^+(X, \mathcal{S})$, prove that*

$$\int g \, d\lambda = \int g \, f \, d\mu$$

where $f = d\lambda/d\mu$ is the Radon - Nikodym derivative of λ with respect to μ.

Exercise 15.6.4. *Let λ, ν and μ be σ - finite measures on $\mathcal{S}$, a σ - algebra of subsets of a set X. Use the previous exercise to show that if $\nu \ll \lambda$ and $\lambda \ll \mu$, then*

$$\frac{d\nu}{d\mu} = \frac{d\nu}{d\lambda}\frac{d\lambda}{d\mu}, \ \mu \text{ a.e.}$$

Further, if λ_1 *and* λ_2 *are absolutely continuous with respect to* μ*, then*

$$d(\lambda_1 + \lambda_2)/d\mu = d\lambda_1/d\mu + d\lambda_2/d\mu \; \mu \; a.e.$$

Exercise 15.6.5. *Prove the results of Comment 15.5.4.*

Chapter 16

Connections To Riemann Integration

Chapter 17

Fubini Type Results

Chapter 18

Interesting Questions

Here are some interesting questions that will probe your understanding of what we have done throughout the course of these notes.

18.1 Midterm Examination

1. This is Exercise 12.5.1.
   Let $X = (0, 1]$. Let $\mathcal{A}$ consist of the empty set and all finite unions of half- open intervals of the form $(a, b]$ from X. Prove $\mathcal{A}$ is an algebra of sets of $(0, 1]$.

2. This is Exercise 12.5.2.
   Let $\mathcal{A}$ be the algebra of subsets of $(0, 1]$ given in Exercise 12.5.1. Let f be an arbitrary function on $[0, 1]$. Define ν_f on $\mathcal{A}$ by

$$\nu_f\Big((a, b]\Big) \;=\; f(b) \;-\; f(a).$$

   Extend ν_f to be additive on finite disjoint intervals as follows: if $(A_i) = (a_i, b_i])$ is a finite collection of disjoint intervals of $(0, 1]$, we define

$$\nu_f\Big(\cup_{i=1}^{n} (a_i, b_i]\Big) \;=\; \sum_{i=1}^{n} f(b_i) \;-\; f(a_i).$$

   (a) Prove that ν_f is additive on $\mathcal{A}$.

   Hint. *It is enough to show that the value of $\nu_f(A)$ is independent of the way in which we write A as a finite disjoint union.*

   (b) Prove ν_f is non negative if and only if f is non decreasing.

3. This is Exercise 12.5.3.
   If λ is an additive set function on an algebra of subsets $\mathcal{A}$, prove that λ can not take on both the value ∞ and $-\infty$.

   Hint. *If there is a set A in the algebra with $\lambda(A) = \infty$ and there is a set B in the algebra with $\lambda(B) = -\infty$, then we can find disjoint sets A' and B' in $\mathcal{A}$ so that $\lambda(A') = \infty$ and $\lambda(B') = -\infty$. But this is not permitted as the value of $\lambda(A' \cup B')$ must be a well - defined extended real value not the undefined value $\infty - \infty$.*

4. This is Exercise 12.5.4.
   Let $\mathscr{T}$ be a covering family for a nonempty set X. Let τ be a non negative, possibly infinite valued premeasure. For any A in X, define

$$\mu^*(A) \quad = \quad \inf \{ \sum_n \tau(T_n) \,|\, T_n \in \mathscr{T}, \, A \subseteq \cup_n T_n \}$$

   where the sequence of sets (T_n) from $\mathscr{T}$ is finite or countably infinite. In the case where there are no sets from $\mathscr{T}$ that cover A, we define the infimum over the resulting empty set to be ∞.

   Prove μ^* is an outer measure on X.

5. This is Exercise 12.5.5.
   Let $X = \{1. 2, 3\}$ and $\mathscr{T}$ consist of $\emptyset$, X and all doubleton subsets $\{x, y\}$ of X. Let τ satisfy

   (i): $\tau(\emptyset) = 0$.

   (ii): $\tau\Big(\{x, y\}\Big) = 1$ for all $x \neq y$ in X.

   (iii): $\tau(X) = 2$.

   (a): Prove the method of Exercise 12.5.4 gives rise to an outer measure μ^* defined by $\mu^*(\emptyset) = 0$, $\mu^*(X) = 2$ and $\mu^*(A) = 1$ for any other subset A of X.

   (b): Now do the construction process again letting $\tau(X) = 3$. What changes?

6. This is Exercise 12.5.6.
   Let X be the natural numbers $\mathbb{N}$ and let τ consist of $\emptyset$, $\mathbb{N}$ and all singleton sets. Define $\tau(\emptyset) = 0$ and $\tau(\{x\}) = 1$ for all x in $\mathbb{N}$.

   (a): Let $\tau(\mathbb{N}) = 2$. Prove the method of Exercise 12.5.4 gives rise to an outer measure μ^*. Determine the family of measurable sets (i.e., the sets that satisfy the Caratheodory Condition).

   (b): Let $\tau(\mathbb{N}) = \infty$ and answer the same questions as in Part (a).

   (c): Let $\tau(\mathbb{N}) = 2$ and set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a).

   (d): Let $\tau(\mathbb{N}) = \infty$ and again set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a). You should see $\mathbb{N}$ is measurable but $\tau(\mathbb{N}) \neq \mu(\mathbb{N})$, where μ denotes the measure constructed in the process of Part (a).

   (e): Let $\tau(\mathbb{N}) = 1$ and again set $\tau(\{x\}) = 2^{-(x-1)}$. Now answer the same questions as in Part (a). What changes?

18.2 Final Examination

1. This is Exercise 11.5.1.
   Let $(X, \mathcal{S}, \mu)$ be a measure space. Let f be in $\mathcal{L}_p(X, \mathcal{S}\mu)$ for $1 \leq p < \infty$. Let $E = \{x \,|\, |f(x)| \neq 0\}$. Prove E is σ - finite.

2. This is Exercise 11.5.2.
   Let $(X, \mathcal{S}, \mu)$ be a finite measure space. If f is measurable, let $E_n = \{x \mid n - 1 \leq |f(x)| < n\}$. Prove f is in $\mathcal{L}_1(X, \mathcal{S}\mu)$ if and only if $\sum_{n=1}^{\infty} n\mu(E_n) < \infty$.

   More generally, prove f is in $\mathcal{L}_p(X, \mathcal{S}\mu)$, $1 \leq p < \infty$, if and only if $\sum_{n=1}^{\infty} n^p \mu(E_n) < \infty$.

3. This is Exercise 14.5.6.
   Let $(\Re, \mathcal{M}, \mu)$ denote the measure space consisting of the Lebesgue measurable sets $\mathcal{M}$ and Lebesgue measure μ. Let the sequence (f_n) of measurable functions be defined by

   $$f_n \quad = \quad n\, I_{[1/n, 2/n]}.$$

   Prove $f_n \to 0$ on all $\Re$, $f_n \to 0\; [meas]$ but $f_n \not\to 0\; [p - norm]$ for $1 \leq p < \infty$.

4. This is Exercise 15.6.1.
   Let $(X, \mathcal{S})$ be a measurable space and λ is a charge on $\mathcal{S}$. Prove if P_1 and P_2 are positive sets for λ, then $P_1 \cup P_2$ is also a positive set for λ.

5. This is Exercise 9.7.5.
   Let $(X, \mathcal{S})$ be a measurable space. Let (μ_n) be a sequence of measures on $\mathcal{S}$ with $\mu_n(X) \leq 1$ for all n. Define λ on $\mathcal{S}$ by

   $$\lambda(E) \quad = \quad \sum_{n=1}^{\infty} 1/2^n\, \mu_n(E)$$

   for all measurable E. Prove λ is a measure on $\mathcal{S}$.

6. This is Exercise 10.8.4.
   Let $(X, \mathcal{S})$ be a measurable space. Let $\mathcal{C}$ be the collection of all charges on $\mathcal{S}$. Prove that $\mathcal{C}$ is a Banach Space under the operations

   $$\begin{aligned} \Big(c\,\mu\Big)(E) &= c\,\mu(E),\ \forall\, c\, \in\, \Re,\ \forall\, \mu \\ \Big(\mu\, +\, \nu\Big)(E) &= \mu(E)\, +\, \nu(E),\ \forall\, \mu,\ \nu, \end{aligned}$$

   with norm $\|\, \mu\, \| = \, |\mu|(X)$

7. This is Exercise 13.4.1.
   A family $\mathcal{A}$ of subsets of the set X is an **algebra** if

   (i): $\emptyset$, X are in $\mathcal{A}$.

   (ii): $E \in \mathcal{A}$ implies $E^C \in \mathcal{A}$.

   (iii): if $\{A_1, \ldots, A_n\}$ is a finite collection of sets in $\mathcal{A}$, then their union is in $\mathcal{A}$.

Further, the mapping τ is sometimes called a pseudo-measure on the algebra $\mathcal{A}$ if $\tau : \mathcal{A} \rightarrow [0, \infty]$ and

(i): $\tau(\emptyset) = 0$.

(ii): If (A_i) is a countable collection of disjoint sets in $\mathcal{A}$ whose union is also in $\mathcal{A}$ (note this is not always true because $\mathcal{A}$ is not a σ - algebra), then

$$\tau(\cup_i A_i) \quad = \quad \sum_i \tau(A_i).$$

Now we get to the exercise:

(a): Let $\mathcal{U}$ be the family of subsets of $\Re$ of the form $(a, b]$, $(-\infty, b]$, (a, ∞) and $(-\infty, \infty)$. Prove $\mathcal{F}$, the collection of all finite unions of sets from $\mathcal{U}$ is an algebra of subsets of $\Re$.

(b): Prove τ equal to the usual length of an interval is a pseudo-measure on $\mathcal{F}$.

(c): Let g be any monotone increasing function on $\Re$ which is continuous from the right. This means

$$\begin{aligned} \lim_{h \rightarrow 0^+} g(x+h) \ \text{ exists }, \ \forall\, x, \\ \lim_{x \rightarrow -\infty} g(x) \ \text{ exists}, \\ \lim_{x \rightarrow \infty} g(x) \ \text{ exists.} \end{aligned}$$

where the last two limits could be $-\infty$ and ∞ respectively. Define the mapping τ_g on $\mathcal{U}$ by

$$\begin{aligned} \tau_g\Big((a, b]\Big) &= g(b) \ - \ g(a), \\ \tau_g\Big((-\infty, b)\Big) &= g(b) \ - \lim_{x \rightarrow -\infty} g(x), \\ \tau_g\Big((a, \infty)\Big) &= \lim_{x \rightarrow \infty} g(x) \ - \ g(a), \\ \tau_g\Big((-\infty, \infty)\Big) &= \lim_{x \rightarrow \infty} g(x) \ - \lim_{x \rightarrow -\infty} g(x). \end{aligned}$$

and extend τ_g to $\mathcal{F}$ as usual. Prove that τ_g is a pseudo-measure on $\mathcal{F}$.

(d): τ_g can then be used to define an outer measure μ_g^* as usual. There is then an associated σ - algebra of μ_g^* measurable sets of $\Re$, $\mathcal{M}_g$, and μ_g^* restricted to $\mathcal{M}_g$ is a measure, μ_g.

We now prove $\mathcal{F}$ is contained in $\mathcal{M}_g$. Here is the hint for any set I from $\mathcal{F}$. Compare this problem to Example 12.4.1 and Example 12.4.2 which are almost identical in spirit (although the g here is more general) even though they are couched in terms of pre-measures instead of pseudo-measures.

Hint. *Let T be any subset of $\Re$. Let $\epsilon > 0$ be given. Then there is a cover (A_n) of sets from the algebra $\mathcal{F}$ so that*

$$\sum_n \tau_g(A_n) \quad \leq \quad \mu_g^*(T) \ + \ \epsilon.$$

Now $I \cap T \subseteq \cup_n (A_n \cap I)$ and $I^C \cap T \subseteq \cup_n (A_n \cap I^C)$. So

$$\mu_g^*(T \cap I) \leq \sum_n \tau_g(A_n \cap I),$$
$$\mu_g^*(T \cap I^C) \leq \sum_n \tau_g(A_n \cap I^C).$$

Combining, and using the additivity of τ_g, we see

$$\mu_g^*(T \cap I) + \mu_g^*(T \cap I^C) \leq \sum_n \tau_g(A_n) \leq \mu_g^*(T) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have shown I satisfies the Caratheodory condition and so in μ_g^* measurable.

Once you have shown these things, we know the Borel σ - algebra $\mathcal{B}$ is contained in $\mathcal{M}_g$! Measures constructed this way are called Borel - Stieljes measures on $\Re$ when we restrict them to $\mathcal{B}$. If we use the full σ - algebra, we call them Lebesgue - Stieljes measures.

8. This is Exercise 13.4.2.
   Let h be our Cantor function

$$h(x) = (x + \Psi(x))/2.$$

   From the previous exercise, we know τ_h defines a Borel - Stieljes measure. Determine if τ_h is absolutely continuous with respect to the Borel measure on $\Re$ (Borel measure is just Lebesgue measure restricted to $\mathcal{B}$.

9. This is Exercise 15.6.2.
   Let $g_1(x) = 2x$, $g_2(x) = I_{[0,\infty)}$, $g_3(x) = x\, I_{[0,\infty)}$ and $g_4(x) = arctan(x)$. All of these functions generate Borel - Stieljes measures on $\Re$.

   (i): Determine which are absolutely continuous with respect to Borel measure. Then, if absolutely continuous with respect to Borel measure, find their Radon - Nikodym derivative.

   (ii): Which of these measures are singular with respect to Borel measure?

10. This is Exercise 15.6.3.
    Let λ and μ be σ - finite measures on $\mathcal{S}$, a σ - algebra of subsets of a set X. Assume λ is absolutely continuous with respect to μ. If $g \in M^+(X, \mathcal{S})$, prove that

$$\int g\, d\lambda = \int g\, f\, d\mu$$

    where $f = d\lambda/d\mu$ is the Radon - Nikodym derivative of λ with respect to μ.

11. This is Exercise 15.6.4.
    Let λ, ν and μ be σ - finite measures on $\mathcal{S}$, a σ - algebra of subsets of a set X. Use the previous exercise to show that if $\nu \ll \lambda$ and $\lambda \ll \mu$, then

$$\frac{d\nu}{d\mu} = \frac{d\nu}{d\lambda}\frac{d\lambda}{d\mu},\ \mu\, a.e.$$

Further, if λ_1 and λ_2 are absolutely continuous with respect to μ, then

$$d(\lambda_1 + \lambda_2)/d\mu = d\lambda_1/d\mu + d\lambda_2/d\mu \; \mu \; a.e.$$

Part III

References

Bibliography

[1] A. Bruckner, J. Bruckner, and B. Thomson. *Real Analysis.* Prentice - Hall, 1997.

[2] S. Douglas. *Introduction To Mathematical Analysis.* Addison-Wesley Publishing Company, 1996.

[3] W. Fulks. *Advanced Calculus: An Introduction to Analysis.* John Wiley & Sons, third edition, 1978.

[4] H. Sagan. *Advanced Calculus of real valued functions of a Real Variable and Vector - Valued Functions of a Vector Variable.* Houghton Mifflin Company, 1974.

[5] G. Simmons. *Introduction to Topology and Modern Analysis.* McGraw-Hill Book Company, 1963.

[6] K. Stromberg. *Introduction To Classical Real Analysis.* Wadsworth International Group and Prindle, Weber and Schmidt, 1981.

[7] A. Taylor. *General Theory of Functions and Integration.* Dover Publications, Inc., 1985.

Part IV

Detailed Indices

Index

Definition
- $RS[g, a, b]$, 110
- Absolute Continuity Of A Measure, 178
- Absolute Continuity Of Charges, 289
- Additive Set Function, 229
- Algebra Of Sets, 216
- Almost Uniform Convergence, 258
- Caratheodory Condition, 216
- Cauchy Sequence In Norm, 199
- Cauchy Sequences In Measure, 259
- Charges, 160
- Common Refinement Of Two Partitions, 32
- Complete Measure, 172
- Complete NLS, 199
- Conjugate Index Pairs, 195
- Content Of Open Interval, 235
- Continuous Almost Everywhere, 95
- Convergence In Measure, 259
- Convergence Pointwise and Pointwise a.e., 258
- Convergence Uniformly, 258
- Darboux Integrability, 62
- Darboux Lower And Upper Integrals, 61
- Darboux Upper and Lower Sums, 58
- Equivalent Conditions For The Measurability of a Function, 146
- Equivalent Conditions For The Measurability of an Extended Real Valued Function, 150
- Essentially Bounded Functions, 206
- Extended Real Number System, 143
- Functions Of Bounded Variation, 44
- Inner Product Space, 212
- Integral Of A Nonnegative Measurable Function, 167
- Integral Of A Simple Function, 167
- Lebesgue Outer Measure, 236
- Limit Inferior And Superior Of Sequences Of Sets, 163
- Measurability of a Function, 145
- Measurability Of Extended Real Valued Functions, 150
- Measures, 160
- Metric On A Set, 192
- Metric Outer Measure, 222
- Monotone Function, 34
  - Associated Saltus Function, 38
- Norm Convergence, 192
- Norm On A Vector Space, 191
- Outer Measure, 215
- Partition, 32
- Positive and Negative Sets For a Charge, 279
- Premeasures and Covering Families, 227

Propositions Holding Almost Everywhere, 166
Pseudo-Measure, 231
Refinement Of A Partition, 32
Regular Outer Measures, 229
Rewriting Lebesgue Outer Measure Using Edge Length Restricted Covers, 246
Riemann - Stieljes Criterion For Integrability, 118
Riemann - Stieljes Darboux Integral, 118
Riemann - Stieljes Sum, 109
Riemann Integrability Of a Bounded f, 55
Riemann Integrability Of A Bounded Function, 15
Riemann Sum, 13, 55
Riemann's Criterion for Integrability, 62
Set of Extended Real Valued Measurable Functions, 150
Sets Of Content Zero, 95
Sigma - Algebra Generated By Collection A, 141
Sigma Algebra, 139
Simple Functions, 166
Singular Measures, 297
Space Of p Summable Functions, 195
Spaces of Essentially Bounded Functions, 206
Step Function, 112
Summable Functions, 180
The Continuous Part Of A Monotone Function, 39
The Discontinuity Set Of A Monotone Function, 36
The Generalized Cantor Set, 103
The Positive and Negative Parts Of a Charge, 280
Upper and Lower Riemann - Stieljes Darboux Sums, 117
Upper and Lower Riemann - Stieljes Integrals, 118
Variation of a Charge, 286

Integration
Antiderivatives Of Simple Powers, 21
Antiderivatives of Simple Trigonometric Functions, 21
Definite Integrals Of Simple Powers, 21
Definite Integrals Of Simple Trigonometric Functions, 22
Functions With Jump Discontinuities, 27
Functions With Removable Discontinuities, 26
Symbol For The Antiderivative of f is $\int f$, 21
Symbol For The Definite Integral of f on $[a, b]$ is $\int_a^b f(t)\,dt$, 21
The indefinite integral of f is also the antiderivative, 21

Lemma
$M_\delta = \mu^*$, 247
$f = g$ on (a, b) Implies Riemann Integrals Match, 84
f Zero On (a, b) Implies Zero Riemann Integral, 83
Outer Measure Of The Closure Of Interval Equals Content Of Interval, 245
Absolute Continuity Of The Integral, 290
Approximate Finite Lebesgue Covers Of $\overline{I}$., 246
Characterizing Limit Inferior And Superiors Of Sequences Of Sets, 163

Condition For Outer Measure To Be Regular, 229
Continuity Of The Integral, 271
Continuous Functions Of Finite Measurable Functions Are Measurable, 154
Continuous Functions Of Measurable Functions Are Measurable, 156
De Morgan's Laws, 140
Disjoint Decompositions Of Unions, 165
Epsilon - Delta Version Of Absolute Continuity Of a Charge, 289
Essentially Bounded Functions Bounded Above By Their Essential Bound a.e, 208
Essentially Bounded Functions That Are Equivalent Have The Same Essential Bound, 207
Extended Valued Measurability In Terms Of The Finite Part Of The Function, 150
Extending τ_g To Additive Is Well - Defined, 250
Finite Jump Step Functions As Integrators, 116
Function f Zero a.e. If and Only If Its Integral Is Zero, 178
Function Measurable If and Only If Positive and Negative Parts Measurable, 149
Hahn Decomposition Characterization of a Charge, 285
Infimum Tolerance Lemma, 33
Lebesgue - Stieljes Outer Premeasure Is a Pseudo-Measure, 252
Limit Inferiors And Superiors Of Monotone Sequences Of Sets, 164
Measure Of Monotonic Sequence Of Sets, 162
Monotonicity, 161
Monotonicity Of The Abstract Integral For Non Negative Functions, 170
One Jump Step Functions As Integrators, 113
Outer Measure Of Interval Equals Content Of Interval, 245
p-Summable Cauchy Sequence Condition I, 270
p-Summable Cauchy Sequence Condition II, 272
p-Summable Functions Have p-Norm Arbitrarily Small Off a Set, 269
p-Summable Inequality, 270
Pointwise Infimums, Supremums, Limit Inferiors and Limit Superiors are Measurable, 153
Products of Measurable Functions Are Measurable, 154
Properties of Extended Valued Measurable Functions, 152
Properties of Measurable Functions, 148
Properties Of Simple Function Integrations, 167
Radon - Nikodym Technical Lemma, 291
Real Number Conjugate Indices Inequality, 195
Sums Over Finite Lebesgue Covers Of $\overline{I}$ Dominate Content Of I, 237
Supremum Tolerance Lemma, 34
The Upper And Lower Darboux Integral Is Additive On Intervals, 71
The Upper And Lower Riemann - Stieljes Darboux Integral Is Additive On Intervals, 120

Measure
Borel, 254
Lebesgue - Stieljes, 254

Measurable Cover, 229
Monotone Function
Continuous at x From Left If and Only If $u(x) = 0$, 36
Continuous at x From right If and Only If $v(x) = 0$, 36
Continuous at x If and Only If $u(x) = v(x) = 0$, 36
Left Hand Jump at x, $u(x)$, 36
Right Hand Jump at x, $v(x)$, 36
Total Jump at x, $u(x) + v(x)$, 36
Monotone Functions
Saltus Function
Properties, 39

Partitions
Gauge or Norm, 33
Proposition
Refinements and Common Refinements, 33

Theorem
$L(f, \boldsymbol{\pi}_1) \leq U(f, \boldsymbol{\pi}_2)$, 61
$L(f, g, \boldsymbol{\pi}_1) \leq U(f, g, \boldsymbol{\pi}_2)$, 117
L_1 Semi-norm, 192
L_p Is A Vector Space, 198
L_p Semi-Norm, 198, 208
$RI[a, b]$ Is A Vector Space and $RI(f; a, b)$ Is A Linear Mapping, 56
V_f and $V_f - f$ Are Monotone For a Function f of Founded Variation, 50
$\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$ Implies $L(f, \boldsymbol{\pi}) \leq L(f, \boldsymbol{\pi'})$ and $U(f, \boldsymbol{\pi}) \geq U(f, \boldsymbol{\pi'})$, 59
$\boldsymbol{\pi} \preceq \boldsymbol{\pi}'$ Implies $L(f, g, \boldsymbol{\pi}) \leq L(f, g, \boldsymbol{\pi'})$ and $U(f, g, \boldsymbol{\pi}) \geq U(f, g, \boldsymbol{\pi'})$, 117
$f \in BV[a, b] \cap C[a, b]$ If and Only If V_f and $V_f - f$ Are Continuous and Increasing, 54
$f \in BV[a, b]$ Is Continuous If and Only If V_f Is Continuous, 52
$f \in RS[g, a, b]$ Implies $f \in RS[V_g, a, b]$ and $f \in RS[V_g - g, a, b]$, 121
f Bounded Variation and g Continuous Implies Riemann - Stieljes Integral Exists, 129
f Bounded and Continuous At All But Finitely Many Points Implies f is Riemann Integrable, 86
f Bounded and Continuous At All But One Point Implies f is Riemann Integrable, 85
f Continuous and g Bounded Variation Implies Riemann - Stieljes Integral Exists, 128
f Continuous and g Riemann Integrable Implies $f \circ g$ is Riemann Integrable, 93
μ^* Measurable Sets Form Algebra, 217
μ^* Measurable Sets Properties, 218
A Function Of Bounded Variation Is The Difference of Two Increasing Functions, 51
A Monotone Function Has A Countable Number of Discontinuities, 35
Abstract Integration Is Additive, 176
Almost Uniform Convergence Implies Convergence In Measure, 266
Alternate Characterization Of Essentially Bounded Functions, 206
Approximation Of Non negative Measurable Functions By Monotone Sequences, 155
Approximation Of The Riemann Integral, 78
Average Value For Riemann Integrals, 77
Bounded Differentiable Implies Bounded Variation, 45

Bounded Variation Implies Riemann Integrable, 71
Cauchy - Schwartz Inequality, 196
Cauchy Fundamental Theorem Of Calculus, 20
Cauchy In Measure Implies A Convergent Subsequence, 259
Cauchy In Measure Implies Completeness, 263
Cauchy Schwartz Inequality: Sequence Spaces, 206
Cauchy's Fundamental Theorem, 77
Conditions For OMI-F Measures, 230
Conditions For OMI-FE Measures, 231
Constructing Measures From Non Negative Measurable Functions, 178
Constructing Outer Measures Via Premeasures, 227
Continuous Implies Riemann Integrable, 69
Convergence Relationships On Finite Measure Space, 275
Convergence Relationships On General Measurable Space, 274
Convergence Relationships With p-Domination, 276
Convergent Subsequences Exist, 277
Egoroff's Theorem, 267
Equality a.e. Can Imply Measurability Even If The Measure Is Not Complete, 173
Equality a.e. Implies Measurability If The Measure Is Complete, 172
Equivalent Absolute Continuity Conditions For Charges, 289
Existence Of The Riemann Integral, 16
Extended Monotone Convergence Theorem, 175, 179
Fatou's Lemma, 176
Functions Of Bounded Variation Always Possess Right and Left Hand Limits, 52
Functions Of Bounded Variation Are Bounded, 45
Functions Of Bounded Variation Are Closed Under Addition, 46
Functions Of Bounded Variation Have Countable Discontinuity Sets, 52
Fundamental Abstract Integration Inequalities, 184
Fundamental Riemann Integral Estimates, 57
Fundamental Riemann Stieljes Integral Estimates, 120
Fundamental Theorem Of Calculus, 18
Hölder's Inequality, 196
Hölder's Inequality: $p = 1$, 211
Hölder's Inequality: Sequence Spaces, 205
Hahn Decomposition Associated With A Charge, 284
Inner Product On The Space of Square Summable Equivalence Classes, 213
Integrals Of Summable Functions Create Charges, 184
Integrand Continuous and Integrator Continuously Differentiable Implies Riemann - Stieljes Integrable, 129
Integrand Riemann Integrable and Integrator Continuously Differentiable Implies Riemann - Stieljes Integrable, 130
Integration By Parts, 80
Inverses Of Functions Of Bounded Variation, 47
Lebesgue Decomposition Theorem, 297
Lebesgue Measure Is Regular, 249

Lebesgue Outer Measure Is Metric Outer Measure, 248
Lebesgue's Criterion For The Riemann Integrability of Bounded Functions, 95
Lebesgue's Dominated Convergence Theorem, 186
Leibnitz's Rule, 82
Levi's Theorem, 182
Limit Interchange Theorem For Riemann - Stieljes Integral, 138
Linearity of the Riemann - Stieljes Integral, 110
Measure Induced By Outer Measure, 220
Measure Induced By Outer Measure Is Complete, 221
Minkowski's Inequality, 197
Minkowski's Inequality: Sequence Spaces, 205
Monotone Convergence Theorem, 174
Monotone Functions
  A Partition Sum Estimate, 34
Monotone Functions Are Of Bounded Variation, 45
Monotone Implies Riemann Integrable, 70
Open Set Characterization Lemma, 144
Open Sets in a Metric Space Are OMI Measurable, 222
Open Sets In Metric Space μ^* Measurable If and Only If μ^* Metric Outer Measure, 226
p-Norm Convergence Implies Convergence in Measure, 265
Pointwise a.e. Convergence Plus Domination Implies p-Norm Convergence, 269
Pointwise Limits of Measurable Functions Are Measurable, 153
Products Of Functions Of Bounded Variation Are Of Bounded Variation, 46
Properties of f_c, 39
Properties Of The Riemann Integral, 66
Properties Of The Riemann Stieljes Integral, 119
Radon - Nikodym Theorem, 293
Representing The Cantor Set, 104
Riemann - Lebesgue Lemma, 95
Riemann - Stieljes Integral, 110
Riemann Stieljes Fundamental Theorem Of Calculus, 125
Riemann Stieljes Integral Is Additive On Subintervals, 124
Riemann Stieljes Integration By Parts, 111
Sequences Of Equivalence Classes in L_p That Converge Are Cauchy, 199
Sequences That Converge In p - Norm Possess Subsequences Converging Pointwise a.e., 204
Space of Square Summable Equivalence Classes Is A Hilbert Space, 213
Substitution In Riemann Integration, 81
Summable Function Equal a.e. To Another Function With Measure Complete Implies The Other Function Is Also Summable, 182
Summable Function Equal a.e. To Another Measurable Function Implies The Other Function Is Also Summable, 181
Summable Function Form A Linear Space, 185
Summable Implies Finite a.e., 181
The Antiderivative of f, 77
The Fundamental Theorem Of Calculus, 74

The Jordan Decomposition Of A Charge, 280
The Mean Value Theorem For Riemann Integrals, 76
The Recapture Theorem, 78
The Riemann Integral Equivalence Theorem, 62
The Riemann Integral Exists On Subintervals, 72
The Riemann Integral Is Additive On Subintervals, 73, 121
The Riemann Integral Is Order Preserving, 58
The Riemann Integral Limit Interchange Theorem, 87
The Riemann Stieljes Integral Equivalence Theorem, 118
The Riemann Stieljes Integral Exists On Subintervals, 120
The Riemann Stieljes Integral Is Order Preserving, 120
The Set Of Equivalence Classes of L_1 Is A Normed Linear Space, 194
The Set Of Equivalence Classes of L_∞ Is A Normed Linear Space, 209
The Set Of Equivalence Classes of L_p Is A Normed Linear Space, 199
The Space of Equivalence Class of L_∞ Is A Banach Space, 210
The Space of Equivalence Class of L_p Is A Banach Space, 200
The Total Variation Is Additive On Intervals, 49
The Total Variation Of A Function Of Bounded Variation, 46
The Upper And Lower Darboux Integral Are Finite, 61
The Upper And Lower Riemann - Stieljes Darboux Integral Are Finite, 117
The Variation Function Of a Function f Of Bounded Variation, 50
Two Riemann Integrable Functions Match At All But Finitely Many Points Implies Integrals Match, 85
Variation of a Charge In Terms of The Plus and Minus Parts, 287
Variation of a Charge is a Measure, 286
Vitali Convergence Theorem, 272
Weierstrass Approximation Theorem, 90
Theorem:f^2, $f_1 f_2$ and $1/f$ Riemann Stieljes Integrable With Respect To g Of Bounded Variation, 124
Theorem:Constant Functions Are Riemann Integrable, 70

Worked Out Solutions
Integration Substitution
$\int (t^2 + 1)\, 2dt$, 22
$\int (t^2 + 1)^3\, 4dt$, 23
$\int \sin(t^2 + 1)\, 5t\, dt$, 25
$\int \sqrt{t^2 + 1}\, 3t\, dt$, 24
$\int_1^5 (t^2 + 2t + 1)^2\, (t + 1)\, dt$, 25