bauldry-fm.qxd 5/28/2009 12:25 PM Page iii

INTRODUCTION TO REAL ANALYSIS

An Educational Approach

William C. Bauldry

Appalachian State University

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

CONTENTS

CHAPTER 3

A BRIEF INTRODUCTION TO LEBESGUE THEORY

Introduction

The span from Newton and Leibniz to Lebesgue covers only 250 years (Figure 3.1). Lebesgue published his dissertation "Intégrale, longueur, aire" ("Integral, length, area") in the *Annali di Matematica* in 1902. Lebesgue developed "measure of a set" in the first chapter and an integral based on his measure in the second.

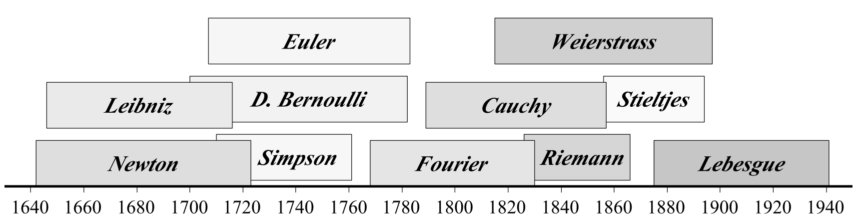


Figure 3.1 From Newton and Leibniz to Lebesgue.

Introduction to Real Analysis. By William C. Bauldry

Copyright © 2009 John Wiley & Sons, Inc.

Part of Lebesgue's motivation were two problems that had arisen with Riemann's integral. First, there were functions for which the integral of the derivative does not recover the original function and others for which the derivative of the integral is not the original. Second, the integral of the limit of a sequence of functions was not necessarily the limit of the integrals. We've seen that uniform convergence allows the interchange of limit and integral, but there are sequences that do not converge uniformly yet the limit of the integrals is equal to the integral of the limit. In Lebesgue's own words from "Integral, length, area" (as quoted by Hochkirchen (2004, p. 272)),

> It thus seems to be natural to search for a definition of the integral which makes integration the inverse operation of differentiation in as large a range as possible.

Lebesgue was able to combine Darboux's work on defining the Riemann integral with Borel's research on the "content" of a set. Darboux was interested in the interplay of the definition of integral with discontinuous functions and in the convergence problems. It was Darboux's development of the Riemann integral that we followed in Chapter 2. Borel (who was Lebesgue's thesis advisor) needed to describe the size of sets of points on which a series converged; he expanded on Jordan's definition of the content of a set which itself was an expansion of Peano's definition of content measuring the size of a set. Peano's work was motivated by Hankel's attempts to describe the size of the set of discontinuities of a Riemann integrable function and by an attempt to define integration analytically, as opposed to geometrically (Hawkins, 2002, chapter 4). Rarely, if ever, is revolutionary mathematics done in isolation. Hochkirchen's (2004) essay "The Theory of Measure and Integration from Riemann to Lebesgue" gives a detailed historical perspective.

Another problem also provided primary motivation for Lebesgue: the question of convergence and integrating series term by term. Newton had used series expansions cleverly to integrate functions when developing calculus. Fourier thought that it was always valid to integrate a trigonometric series representation of a function term by term. Cauchy believed continuity of the terms sufficed; Cauchy's integral required continuity to exist. Then Abel gave an example that didn't work. Weierstrass recognized that uniform convergence was the key to term-by-term integration. Dirichlet developed wildly discontinuous counterexamples such as his "monster." Riemann defined his integral so as to not require continuity, but uniform convergence of the series was still necessary for term-by-term integration. However, some nonuniformly convergent series could still be integrated term by term. What is the right condition? Lebesgue's theory can answer these questions.

We now turn to Lebesgue's concept of the measure of a set.

3.1 LEBESGUE MEASURE AND MEASURABLE SETS

Lebesgue's measure is an extended real-valued set function, a function from a collection of sets into $[0, \infty]$. Measure is based on the lengths of open intervals as these intervals are the basic building blocks of open sets in the reals. The best measure μ would satisfy four properties:

1. For each interval I, $\mu(I) = \text{length}(I)$.
2. For $x \in \mathbb{R}$ and $E \subset \mathbb{R}$, $\mu(x + E) = \mu(E)$.
3. For a sequence of disjoint sets $\{E_n\}$, $\mu\left(\bigcup_n E_n\right) = \sum_n \mu(E_n)$.
4. Every subset of $\mathbb{R}$ can be measured by μ.

Unfortunately, we're asking too much; it's not possible to satisfy all four properties. Property 4 is the first to fall. There will be sets that cannot be measured. (Look ahead to "Vitali's Nonmeasurable Set," p. 249.) If we assume the *continuum hypothesis* (see p. 172) is true, then we cannot have a measure satisfying properties 1, 3, and 4. We are unwilling to give up the first two; we'll weaken the third initially, hoping to recapture it later.

Algebras and σ-Algebras of Sets

An *algebra of sets* is a nonempty collection that behaves well with respect to unions and complements. A collection $\mathcal{A}$ of sets is an *algebra* if and only if the union of any two sets in $\mathcal{A}$ is also in $\mathcal{A}$, and the complement of every set in $\mathcal{A}$ is also in $\mathcal{A}$. De Morgan's law then implies the intersection of any two sets in $\mathcal{A}$ is in $\mathcal{A}$. By extension, any union or intersection of a finite number of sets in $\mathcal{A}$ is also in $\mathcal{A}$. De Morgan's laws give us a principle of duality for statements about sets in an algebra: replace unions with intersections and vice versa to obtain a new true statement.

Suppose that $A \neq \emptyset$ is a member of an algebra $\mathcal{A}$ of subsets of X. Then $A^c \in \mathcal{A}$. Since A and A^c are in $\mathcal{A}$, we have that $A \cup A^c = X$ and $A \cap A^c = \emptyset$ are both members of $\mathcal{A}$. In particular, every algebra of sets of reals contains both $\emptyset$ and $\mathbb{R}$.

An algebra of sets $\mathcal{A}$ is a σ*-algebra* if every union of a countable collection of sets from $\mathcal{A}$ is also in $\mathcal{A}$. Once again, De Morgan's laws tell us that σ-algebras are closed with respect to countable intersections.

■ **EXAMPLE 3.1**

1. For any $X \neq \emptyset$, the collection $\{\emptyset, X\}$ forms a σ-algebra, the *trivial* σ*-algebra*. *Prove this!*
2. Let $\mathcal{P}(\mathbb{N})$ be the *power set* of $\mathbb{N}$, the set of all subsets of $\mathbb{N}$. Then $\mathcal{P}(\mathbb{N})$ is a σ-algebra. *Show this!*
3. Let $\mathcal{F}$ be the collection of subsets of an infinite set X that are finite or have finite complement, the *co-finite algebra*.

   (a) $\mathcal{F}$ is an algebra.

   The complement of any set in $\mathcal{F}$ is clearly in $\mathcal{F}$. *Why?*

   Since the union of two finite sets is finite, the union of two sets with finite complement has finite complement, and the union of a finite set with a set having finite complement has finite complement, $\mathcal{F}$ is closed under unions.

Hence $\mathcal{F}$ is an algebra.

(b) $\mathcal{F}$ is *not* a σ-algebra.

Since X is infinite, we can choose a subset $Y \subset X$ such that Y is countable and Y^c is infinite. For each $y \in Y$, the set $\{y\}$ is finite, and hence in $\mathcal{F}$. However, the union $\bigcup_{y\in Y}\{y\} = Y$, which is not a member of $\mathcal{F}$. *Give an example of such a Y when $X = \mathbb{N}$.*

4. The *Borel σ-algebra on* $\mathbb{R}$ is the smallest σ-algebra containing all the open sets and is denoted by $\mathcal{B}(\mathbb{R})$.

■

Lebesgue Outer Measure

A collection of open intervals $\{I_n \mid n = 1, 2, \dots\}$ *covers* a set E if $E \subseteq \bigcup_n I_n$. Since the intervals are open, we call $\{I_n\}$ an *open cover* of E. Define the *length* ℓ of the open interval $I = (a, b)$ to be $\ell(I) = b - a$. We combine open covers and length to measure the size of a set. Since the cover contains the set, we'll call it the *outer measure*. The outer measure is extremely close to the measure Jordan defined in 1892.

Definition 3.1 (Lebesgue Outer Measure) *For any set $E \subseteq \mathbb{R}$, define the* Lebesgue outer measure μ^* *of E to be*

$$\mu^*(E) = \inf_{E\subset\bigcup I_n} \sum_n \ell(I_n)$$

the infimum of the sums of the lengths of open covers of E.

If $A \subseteq B$, then any open cover of B also covers A. Therefore $\mu^*(A) \le \mu^*(B)$. This property is called *monotonicity*.

Theorem 3.1 *The outer measure of an interval is its length.*

Proof: First, consider a bounded, closed interval $I = [a, b]$. For any $\epsilon > 0$,

$$[a, b] \subset (a - \epsilon/2, b + \epsilon/2)$$

Hence, $\mu^*(I) \le b - a + \epsilon$. Since $\epsilon > 0$ is arbitrary, $\mu^*([a, b]) \le b - a$.

Now let $\{I_n\}$ be an open cover of $[a, b]$. The Heine-Borel theorem states that since $[a, b]$ is closed and bounded there is a finite subcover $\{I_k \mid k = 1..N\}$ for I. Order the intervals so they overlap, starting with the first containing a and ending with the last containing b. *How do we know we can do this?* Thus

$$\begin{aligned}\sum_{k=1}^{N} \ell(I_k) &= (b_1 - a_1) + (b_2 - a_2) + \cdots + (b_N - a_N) \\ &= b_N - (a_N - b_{N-1}) - (a_{N-1} - b_{N-2}) - \cdots - (a_2 - b_1) - a_1 \\ &\geq b_N - a_1 \\ &> b - a\end{aligned}$$

Thus $\mu^*(I) \geq b - a$, which combines with the first inequality to yield $\mu^*(I) = b - a$.

Second, let I be any bounded interval and let $\epsilon > 0$. There is a closed interval $J \subset I$ such that $\ell(I) - \epsilon < \ell(J)$. Then

$$\ell(I) - \epsilon < \ell(J) = \mu^*(J) \leq \mu^*(I) \leq \mu^*(\bar{I}) = \ell(\bar{I}) = \ell(I)$$

or

$$\ell(I) - \epsilon < \mu^*(I) \leq \ell(I).$$

Since $\epsilon > 0$ is arbitrary, we have $\mu^*(I) = \ell(I)$.

Last, if I is an infinite interval, for each $n \in \mathbb{N}$, there is a closed interval $J \subset I$ with $\mu^*(J) = n$. Then $n = \mu^*(J) \leq \mu^*(I)$ implies that $\mu^*(I) = \infty$. ∎

While we can't satisfy additivity (property 3), we can weaken the property by changing the equality to less than or equal. This relation is called *subadditivity*.

Theorem 3.2 *Lebesgue outer measure is* countably subadditive. *Let $\{E_n\}$ be a countable sequence of subsets of $\mathbb{R}$. Then*

$$\mu^*\left(\bigcup_n E_n\right) \leq \sum_n \mu^*(E_n)$$

Proof: If any set E_n has infinite outer measure, the inequality is trivially true. Suppose that $\mu^*(E_n) < \infty$ for all n and let $\epsilon > 0$. For each n, there is a countable collection of open intervals $\{I_{n,j} \mid j \in \mathbb{N}\}$ that covers E_n and such that

$$\sum_{j \in \mathbb{N}} \ell(I_{n,j}) < \mu^*(E_n) + \frac{\epsilon}{2^n}$$

A countable collection of countable sets is countable, so $\{I_{n,j} \mid n, j \in \mathbb{N}\}$ is a countable, open cover of $E = \bigcup E_n$. Therefore

$$\begin{aligned}\mu^*\left(\bigcup_n E_n\right) &\leq \sum_{n,j} \ell(I_{n,j}) = \sum_{n \in \mathbb{N}} \left(\sum_{j \in \mathbb{N}} \ell(I_{n,j})\right) \\ &< \sum_{n \in \mathbb{N}} \left(\mu^*(E_n) + \frac{\epsilon}{2^n}\right) \\ &= \sum_{n \in \mathbb{N}} \mu^*(E_n) + \epsilon\end{aligned}$$

Thus $\mu^*\left(\bigcup_n E_n\right) \leq \sum_n \mu^*(E_n)$. ∎

Several results are immediate corollaries.

Corollary 3.3 *The outer measure of a countable set is zero.*

We know that $\mu^*([0,1]) = 1$.

Corollary 3.4 *The interval* $[0,1]$ *is not countable.*

The union of a set of open intervals is an open set. The definition of infimum tells us that there is an open cover with outer measure within ϵ of the set it covers. Combine these observations to find an open set with outer measure arbitrarily close to the outer measure of a given set.

Corollary 3.5 *Let* $E \subseteq \mathbb{R}$ *and* $\epsilon > 0$. *There is an open set* O *such that* $E \subseteq O$ *and* $\mu^*(E) \leq \mu^*(O) \leq \mu^*(E) + \epsilon$.

With the outer measure in hand, we ask, "What is a measurable set?"

Lebesgue Measure

In 1902, Lebesgue defined the *inner measure* of a set in terms of the outer measure of the complement of the set. If the outer and inner measures were equal, Lebesgue defined the measure of the set to be that common value. One drawback to this approach is that it can be difficult unless the measure of the whole space is finite. [See, e.g., Hochkirchen (2004).] As a way to generalize the concept, Carathéodory developed a method of constructing a measure using only the outer measure. We'll follow Carathéodory's lead. First, we determine which sets are measurable, then we define the Lebesgue measure for these sets.

Definition 3.2 (Carathéodory's Condition) *A set* E *is* Lebesgue measurable *if and only if for every set* $A \subseteq \mathbb{R}$ *we have*

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

Let $\mathfrak{M}$ *be the family of all Lebesgue measurable sets.*

Informally, a set is measurable if it splits every other set into two pieces with measures that add correctly. The definition of measurable is symmetric: if E is measurable, so is E^c; i.e., if $E \in \mathfrak{M}$, then $E^c \in \mathfrak{M}$. Also, it is easily seen that $\emptyset$ and $\mathbb{R} \in \mathfrak{M}$.

Theorem 3.6 *If* $\mu^*(E) = 0$, *then* E *is measurable.*

Proof: For any set A, it's true that

$$\mu^*(A) = \mu^*((A \cap E) \cup (A \cap E^c)) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Since $A \cap E \subseteq E$, we see that $\mu^*(A \cap E) \leq \mu^*(E) = 0$. Thus $\mu^*(A \cap E) = 0$. Now note that $A \cap E^c \subseteq A$, so $\mu^*(A \cap E^c) \leq \mu^*(A)$. Hence

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c) \leq \mu^*(A)$$

Thus $E \in \mathfrak{M}$. ∎

If two sets are measurable, is their union also measurable?

Theorem 3.7 *The union of two measurable sets is measurable.*

Proof: Let E_1 and E_2 be two measurable sets. Let A be a set to use in Carathédory's condition. Apply Carathédory's condition to the set $A \cap E_1^c$ to have

$$\mu^*(A \cap E_1^c) = \mu^*((A \cap E_1^c) \cap E_2) + \mu^*((A \cap E_1^c) \cap E_2^c)$$

Now

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

so

$$\begin{aligned} &\mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1 \cup E_2)^c) \\ &\quad = \mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1^c \cap E_2^c)) \\ &\quad \le \left[\mu^*(A \cap E_1) + \mu^*(A \cap E_2 \cap E_1^c)\right] + \mu^*(A \cap E_1^c \cap E_2^c) \\ &\quad = \mu^*(A \cap E_1) + \left[\mu^*(A \cap E_2 \cap E_1^c) + \mu^*(A \cap E_1^c \cap E_2^c)\right] \\ &\quad = \mu^*(A \cap E_1) + \left[\mu^*((A \cap E_1^c) \cap E_2) + \mu^*((A \cap E_1^c) \cap E_2^c)\right] \end{aligned}$$

Since E_2 is measurable, then

$$\mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1 \cup E_2)^c) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c)$$

Now, since E_1 is measurable, we have

$$\mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1 \cup E_2)^c) = \mu^*(A)$$

Therefore

$$\mu^*(A) = \mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap (E_1 \cup E_2)^c)$$

which yields $E_1 \cup E_2 \in \mathfrak{M}$. ∎

The results above indicate that the collection of measurable sets $\mathfrak{M}$ forms an algebra.

Theorem 3.8 *A countable union of measurable sets is measurable.*

Proof: Let $\{E_k\}$ be a countable sequence of measurable sets and put $E = \bigcup E_k$. Let A be an abitrary test set to use in Carathédory's condition. By subadditivity,

$$\mu^*(A) \le \mu^*((A \cap E) + \mu^*(A \cap E^c)$$

Set $F_n = \bigcup_{k=1}^n E_k$ and $F = \bigcup_{k=1}^\infty E_k = E$. Also, set $G_1 = E_1$, $G_2 = E_2 - E_1$, $G_3 = E_3 - E_1 - E_2$, etc., and $G = \bigcup_{k=1}^\infty G_k$. Then

1. $G_i \cap G_j = \emptyset$ for $i \neq j$ (G_k *are pairwise disjoint*)
2. $F_n = \displaystyle\bigcup_{k=1}^{n} G_k$
3. $G = F = E$

Each F_n and G_n are measurable. *Verify this!* Therefore,

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

Apply Carathédory's condition to $A \cap F_n$ to see

$$\begin{aligned}\mu^*(A \cap F_n) &= \mu^*((A \cap F_n) \cap G_n) + \mu^*((A \cap F_n) \cap G_n^c) \\ &= \mu^*(A \cap G_n) + \mu^*(A \cap F_{n-1})\end{aligned}$$

Why? Iterate this relation to obtain

$$\mu^*(A \cap F_n) = \sum_{k=1}^{n} \mu^*(A \cap G_k)$$

Since $F_n \subseteq F$, then $F^c \subseteq F_n^c$ for each n, so $\mu^*(A \cap F_n^c) \geq \mu^*(A \cap F^c)$. Hence

$$\mu^*(A) \geq \sum_{k=1}^{n} \mu^*(A \cap G_k) + \mu^*(A \cap F^c)$$

The summation above is increasing and bounded (*Why?*), and therefore convergent. So

$$\mu^*(A) \geq \sum_{k=1}^{\infty} \mu^*(A \cap G_k) + \mu^*(A \cap F^c)$$

But

$$\sum_{k=1}^{\infty} \mu^*(A \cap G_k) \geq \mu^* \left(\bigcup_{k=1}^{\infty} (A \cap G_k) \right) = \mu^* \left(A \cap \bigcup_{k=1}^{\infty} G_k \right) = \mu^*(A \cap F)$$

So we have

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

Since we have already shown the reverse inequality, then

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

Therefore, by Carathédory's condition, $F = \bigcup E_k \in \mathfrak{M}$. ■

We have shown:

Corollary 3.9 *The collection of measurable sets $\mathfrak{M}$ is a σ-algebra.*

Corollary 3.10 *The Borel sets are measurable.*

Proof: Since $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra containing all the open sets and $\mathfrak{M}$ is a σ-algebra containing all open intervals, and hence all open sets, then $\mathcal{B}(\mathbb{R}) \subseteq \mathfrak{M}$. ■

There are measurable sets that are not Borel sets. The construction of a measurable, non-Borel set is beyond our scope; see Cohn (1980, p. 56) for details.

Even though $\mathfrak{M}$ is huge, there must be sets that are not measurable. So $\mathfrak{M} \subsetneqq \mathcal{P}(\mathbb{R})$. Vitali (1905) constructed the first example of a nonmeasurable set. (Look ahead to "Vitali's Nonmeasurable Set," p. 249.) Vitali's construction leads to the result that any translation-invariant measure (Exercise 3.8) has nonmeasurable sets.

Now that we've studied the class of measurable sets $\mathfrak{M}$, we are ready to define Lebesgue measure.

Definition 3.3 (Lebesgue Measure) *The* Lebesgue measure μ *is the restriction of the outer measure μ^* to the measurable sets $\mathfrak{M}$. That is, for $E \in \mathfrak{M}$, set $\mu(E) = \mu^*(E)$.*

■ **EXAMPLE 3.2**

Any set E with outer measure zero is measurable and so is Lebesgue measurable with $\mu(E) = 0$.

Definition 3.4 (Almost Everywhere) *A property is said to hold* almost everywhere *if and only if the measure of the set on which the property does not hold is zero.*

■ **EXAMPLE 3.3**

Let $\chi_{\mathbb{N}}(x)$ be the characteristic function of the integers, that is,

$$\chi_{\mathbb{N}}(x) = \begin{cases} 1 & x \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

Since $\mu(\mathbb{N}) = 0$ (*Show this!*), then $\chi_{\mathbb{N}} = 0$ almost everywhere. ■

The term almost everywhere is often abbreviated "a.e." Is $\chi_{\mathbb{N}}$ continuous a.e.?

Outer measure is countably subadditive; therefore Lebesgue measure is countably subadditive, too. However, if we have a sequence of pairwise disjoint sets, we can say more. We can recapture property 3, additivity!

Theorem 3.11 (Additivity of Lebesgue Measure) *Let $\{E_n\}$ be a countable sequence of pairwise disjoint, measurable sets. Then*

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

Proof: If $\{E_n\}$ is a finite sequence, we'll use induction on n, the number of sets in the sequence. For $n = 1$, the result trivially holds. Assume the equation holds for $n-1$ sets. Since the sets are disjoint,

$$\left(\bigcup_{k=1}^{n} E_k\right) \cap E_n = E_n \quad \text{and} \quad \left(\bigcup_{k=1}^{n} E_k\right) \cap E_n^c = \bigcup_{k=1}^{n-1} E_k$$

A finite union of measurable sets is measurable and E_n is measurable; thus

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{n} E_k\right) &= \mu\left(\left[\bigcup_{k=1}^{n} E_k\right] \cap E_n\right) + \mu\left(\left[\bigcup_{k=1}^{n} E_k\right] \cap E_n^c\right) \\
&= \mu\left(E_n\right) + \mu\left(\bigcup_{k=1}^{n-1} E_k\right) \\
&= \mu\left(E_n\right) + \sum_{k=1}^{n-1} \mu(E_k)
\end{aligned}$$

The result follows by induction.

If $\{E_n\}$ is an infinite sequence, then $\bigcup_{k=1}^{n} E_k \subset \bigcup_{k=1}^{\infty} E_k$. Hence

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \geq \mu\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mu(E_k)$$

We have a bounded, increasing series on the right. Thus

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \geq \sum_{k=1}^{\infty} \mu(E_k)$$

Countable subadditivity supplies the reverse inequality, so

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

Hence the result holds for finite or countably infinite sequences of disjoint sets. ∎

■ EXAMPLE 3.4

Let $a \in \mathbb{R}$ and set $I = (a, \infty)$. Show $\mu(I) = \infty$.

We already know that $\mu((a, \infty)) = \infty$ is true, but we'll use additivity for a different verification. Define $E_k = (a + k - 1, a + k]$ for $k = 1, 2, \ldots$. Then $\mu(E_k) = 1$ for all k. Further, $E_i \cap E_j = \emptyset$ for $i \neq j$; that is, the sequence is pairwise disjoint. So

$$\mu((a, \infty)) = \mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k) = \sum_{k=1}^{\infty} 1 = \infty$$

∎

If we have a decreasing sequence of nested sets, then we can calculate the measure of the intersection as the limit of the measures. In a certain sense, this statement says "the measure of the limit is the limit of measures" for nested sets. This result will be quite useful to us in integration.

Theorem 3.12 *Let $\{E_n\}$ be an infinite sequence of nested, measurable sets, that is, for each n, $E_{n+1} \subset E_n$. If $\mu(E_1)$ is finite, then*

$$\mu\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{n\to\infty} \mu(E_n)$$

Proof: Let $E = \bigcap_k E_k$ and define $F_k = E_k - E_{k+1}$. The F_k are pairwise disjoint because the E_k are nested, and

$$\bigcup_{k=1}^{\infty} F_k = E_1 - E$$

Since E and E_1 are measurable, so is $E_1 - E$. Then

$$\mu(E_1 - E) = \mu\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mu(F_k)$$

By the definition of F_k, we have

$$\mu(E_1 - E) = \sum_{k=1}^{\infty} \mu(E_k - E_{k+1})$$

Since $E_1 \supset E_2 \supset E_3 \supset \cdots \supset E$ and $\mu(E_1) < \infty$, the measure of all E_k and of E is finite. Thus, $\mu(E_1) = \mu(E) + \mu(E_1 - E)$ (*Why?*) implies that $\mu(E_1 - E) = \mu(E_1) - \mu(E)$. Continuing, $\mu(E_k) = \mu(E_{k+1}) + \mu(E_k - E_{k+1})$ implies that $\mu(E_k - E_{k+1}) = \mu(E_k) - \mu(E_{k+1})$ for each k. Substituting, we see

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

Since $\mu(E_1) < \infty$, the result holds. ∎

■ **EXAMPLE 3.5 The Cantor Set**

We will define the *Cantor set* C as the infinite intersection of a nested sequence of sets. Set $C_0 = [0,1]$ and note that $\mu(C_0) = 1$. Now define C_1 to be C_0 minus the open middle third interval; $C_1 = [0,1/3] \cup [2/3,1]$. Note $\mu(C_1) = 2/3$. Continue the process, removing the middle third of each interval, to have

$$C_2 = \left[0, \frac{1}{9}\right] \bigcup \left[\frac{2}{9}, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, \frac{7}{9}\right] \bigcup \left[\frac{7}{9}, 1\right]$$
$$C_3 = \left[0, \frac{1}{27}\right] \bigcup \left[\frac{2}{27}, \frac{3}{27}\right] \bigcup \left[\frac{6}{27}, \frac{7}{27}\right] \bigcup \left[\frac{8}{27}, \frac{9}{27}\right] \bigcup \left[\frac{18}{27}, \frac{19}{27}\right] \bigcup \left[\frac{20}{27}, \frac{21}{27}\right] \bigcup \left[\frac{24}{27}, \frac{25}{27}\right] \bigcup \left[\frac{26}{27}, 1\right]$$

and so forth. Figure 3.2 shows several iterations of the procedure.

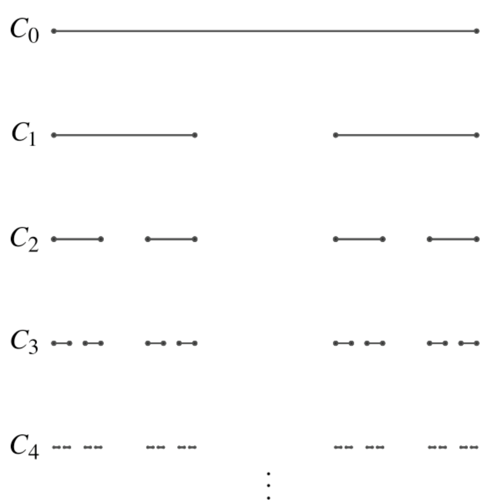


Figure 3.2 Constructing the Cantor Set

Each C_n is a finite union of closed intervals, and thus closed. For every n, $C_{n+1} \subset C_n$, and the measure of C_{n+1} is $2/3$ the measure of C_n. Hence

$$\begin{aligned}
\mu(C_0) &= 1 \\
\mu(C_1) &= \frac{2}{3}\,C_0 = \frac{2}{3} \\
\mu(C_2) &= \frac{2}{3}\,C_1 = \left(\frac{2}{3}\right)^2 \\
&\vdots \\
\mu(C_n) &= \frac{2}{3}\,C_{n-1} = \left(\frac{2}{3}\right)^n \\
&\vdots
\end{aligned}$$

The Cantor set C is the intersection of the C_n's,

$$C = \bigcap_{n=0}^{\infty} C_n$$

Since we are removing the middles of the intervals, the endpoints remain in the sets, doubling the number of endpoints in each iteration. This doubling can be used to create a bijective map from C onto $\mathbb{R}$ showing that C is uncountable.

What is the measure of the Cantor set? By Theorem 3.12, we see

$$\mu(C) = \lim_{n\to\infty} \mu(C_n) = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0$$

The Cantor set is very interesting. The set C is a *perfect set* where every point is an accumulation point but the Cantor set contains no interval. Most fascinating for us, though, is the fact that the Cantor set is an uncountable set with measure zero! ■

The property of being measurable is more general than being open or closed. Consider the interval $[0, 1)$. This interval is neither open nor closed but is easily seen to be measurable. How strange can a measurable set be? In one sense, not very.

Theorem 3.13 *Let $E \subseteq \mathbb{R}$. The following are equivalent.*

1. $E \in \mathfrak{M}$; *i.e., E is measurable,*
2. *To any $\epsilon > 0$, there is an open set $O \supset E$ such that $\mu^*(O - E) < \epsilon$,*
3. *To any $\epsilon > 0$, there is a close set $F \subset E$ such that $\mu^*(E - F) < \epsilon$.*

The proof is left to the exercises.

Define the *symmetric difference* $\triangle$ of two sets A and B by

$$A \triangle B = (A - B) \cup (B - A)$$

Theorem 3.14 *If $E \subset [a, b]$ is measurable, then, for any $\epsilon > 0$, there is a finite collection of open intervals $I_1, I_2, \ldots, I_n$ such that $\mu(E \triangle \bigcup I_n) < \epsilon$.*

Proof: Since E is measurable and any open set in $\mathbb{R}$ is the countable union of open intervals, there is a countable union of open intervals $U = \bigcup I_j$ such that

$$\mu(E) \leq \mu(U) \leq \mu(E) + \frac{\epsilon}{2}$$

Thus $\mu(U - E) < \epsilon/2$. *Why?*

Put $F = [a, b] - E$; then F is measurable. Similarly, there is a countable collection of open intervals V containing F with

$$\mu(V - F) < \frac{\epsilon}{2}$$

Since V is open, V^c is closed. Also,

$$V - F = V - ([a, b] - E) \supset V \cap E = E - V^c$$

Therefore,

$$\mu(E - V^c) \leq \mu(V - F) < \frac{\epsilon}{2}$$

Now, U is a countable collection of open intervals that covers the closed and bounded set $V^c \cap [a, b] \subseteq E$. By the Heine-Borel theorem, U has a finite subcover $W = \bigcup_{j=1}^{n} I_j$ that covers $V^c \cap [a, b]$. Then

$$\begin{aligned} \mu^*(E \triangle W) &\leq \mu^*(E \cap W^c) + \mu^*(E^c \cap W) \\ &\leq \mu^*(E - V^c) + \mu^*(U - E) < \epsilon \end{aligned}$$

(*Why?*), and the result holds. ∎

So a measurable set is "almost" an open set or a closed set. On the other hand, we know there are measurable sets that are not in $\mathcal{B}(\mathbb{R})$, that is, that are not Borel sets. Since $\mathcal{B}(\mathbb{R})$ is a σ-algebra that contains all the open sets, it also contains all countable unions of open sets (or closed sets), and so all countable intersections of countable unions of open sets, and so all countable unions of countable intersections of countable unions of open sets, and so forth. A measurable set that is not a Borel must have a very complicated structure. Nevertheless, a measurable set must be sandwiched between a closed set and an open set that are arbitrarily close to each other.

3.2 THE LEBESGUE INTEGRAL

At the turn of the twentieth century, Lebesgue was working on his doctoral dissertation while a professor at the Lycée Centrale, a girls' secondary school, in Nancy, France, from 1899 to 1901. Lebesgue had studied the works of Jordan, Borel, and Baire, and was building on their foundation; he also studied Riemann integration and Fourier series. The problem of certain derivatives not being integrable—functions failing the fundamental theorem—was Lebesgue's motivation. Dirichlet's characteristic function of the rationals (1829) gave an example that could not be Riemann integrated; but his function was nowhere continuous. Thomae's function (1875) is Riemann integrable and is continuous only on the irrationals. Riemann himself had given an example of an integrable function that was discontinuous at any rational whose denominator was divisible by 2 in his 1854 *Habilitation* thesis. Where is the demarcation between integrable and nonintegrable in terms of continuity? Lebesgue's theory answers these questions.

We'll start by classifying the functions to study, then define the Lebegue integral and investigate its properties.

Measurable Functions

Continuity is essentially a local condition describing the behavior of a function at a point. Uniform continuity extends the concept to sets, but still the focus is at the level of a point. We need a more global descriptor. In topology, continuity is described as the inverse image of an open set is an open set. We'll use a similar criterion to define a function as *measurable* in terms of measurable sets.

Theorem 3.15 (Measurability Condition for Functions) *Let f be an extended real-valued function that has a measurable domain D. The following are equivalent.*

1. *For each $r \in \mathbb{R}$, the set $\{x \in D \mid f(x) > r\} = f^{-1}((r, \infty))$ is measurable.*
2. *For each $r \in \mathbb{R}$, the set $\{x \in D \mid f(x) \geq r\} = f^{-1}([r, \infty))$ is measurable.*
3. *For each $r \in \mathbb{R}$, the set $\{x \in D \mid f(x) < r\} = f^{-1}((-\infty, r))$ is measurable.*
4. *For each $r \in \mathbb{R}$, the set $\{x \in D \mid f(x) \leq r\} = f^{-1}((-\infty, r]))$ is measurable.*

Proof: Let $D = \mathrm{dom}(f)$ be a measurable set.

Then $1 \Rightarrow 2$, since $\{x \mid f(x) \geq r\} = \bigcap_n \{x \mid f(x) > r - 1/n\}$, and the countable intersection of measurable sets is measurable.

Now $2 \Rightarrow 3$, because $\{x \mid f(x) < r\} = D - \{x \mid f(x) \geq r\}$, and the difference of two measurable sets is measurable.

Next, $3 \Rightarrow 4$, as $\{x \mid f(x) \leq r\} = \bigcap_n \{x \mid f(x) < r + 1/n\}$, and again, the countable intersection of measurable sets is measurable.

Last, $4 \Rightarrow 1$, since $\{x \mid f(x) > r\} = D - \{x \mid f(x) \leq r\}$, and, once more, the difference of two measurable sets is measurable. ■

Figure 3.3 shows a function f and the set $A = \{x \mid f(x) > 2.5\}$; note the three components of A.

Corollary 3.16 *If f satisfies any of the measurability conditions, then for each $r \in \mathbb{R}$, the set $\{x \mid f(x) = r\}$ is measurable.*

The converse of the corollary is not true; even if the sets $\{x \mid f(x) = r\}$ are all measurable, the function need not satisfy the measurability conditions.

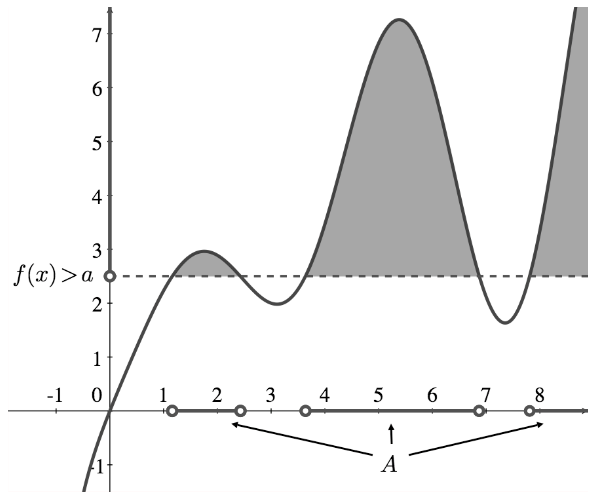


Figure 3.3 The Set $A = \{x \,|\, f(x) > 2.5\}$

Definition 3.5 (Lebesgue Measurable Function) *If an extended real-valued function f has a measurable domain and satisfies the measurability conditions, then f is called a* Lebesgue measurable *or a* measurable function*.*

■ EXAMPLE 3.6 Step Functions

A function $\phi : [a, b] \to \mathbb{R}$ is a *step function* if there is a partition $a = x_0 < x_1 < \cdots < x_n = b$ such that ϕ is constant on each interval $I_k = (x_{k-1}, x_k)$, then

$$\phi(x) = \sum_{k=1}^{n} a_k \chi_{I_k}(x)$$

where $a_k = \phi(I_k)$. See Figure 3.4. (We have not specified the values at x_k, a set of measure zero.) Since the set $\{x \,|\, f(x) \le r\}$ is a finite union of intervals, it is measurable. Thus, step functions are measurable. ■

■ EXAMPLE 3.7 Simple Functions

A function $\psi : [a, b] \to \mathbb{R}$ is a *simple function* if the range of ψ is a finite set $\{a_1, a_2, \ldots, a_n\}$ and, for each $k = 1, \ldots, n$, $\psi^{-1}(a_k)$ is a measurable set. Since

$$\{x \,|\, \psi(x) > r\} = \bigcup_{a_k > r} \psi^{-1}(a_k)$$

and each $\psi^{-1}(a_k)$ is measurable, then ψ is a measurable function.

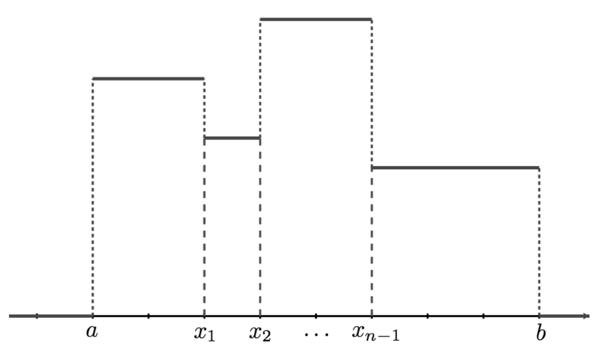


Figure 3.4 A Step Function

An alternate definition is ψ is a simple function if it can be written as

$$\psi(x) = a_1\chi_{E_1}(x) + a_2\chi_{E_2}(x) + \cdots + a_n\chi_{E_n}(x)$$

where $E_1, E_2, \ldots, E_n$ are pairwise-disjoint, measurable sets and $a_1, a_2, \ldots, a_n$ are constants.

1. The characteristic function of the rationals $\chi_{\mathbb{Q}}$ is a simple function, and thus measurable. *Would the characteristic function of a general measurable set be a simple function?*

2. Fix a value of $n \in \mathbb{N}$. Define $\psi_n : [0, n) \to \mathbb{R}$ for rational r by

$$\psi_n(r) = \sum_{k=1}^{n} k \cdot \chi_{[k-1,k)}(r)$$

and $\psi_n(x) = 0$ for irrational x. Then ψ_n is a simple function. *What does ψ_n's graph look like?*

■

Every step function is a simple function, but not conversely.

For a continuous function f with a measurable domain, the inverse image of an open set is open. Open sets are measurable. Hence, continuous functions with measurable domains are measurable functions.

Is the combination of measurable functions still measurable?

Theorem 3.17 (Algebra of Measurable Functions) *Let f and g be measurable functions on a common domain, and let $c \in \mathbb{R}$. Then $f + c$, cf, $f \pm g$, f^2, and fg are measurable.*

Proof: Since $\{x \mid f(x) + c < r\} = \{x \mid f(x) < r - c\}$ and, for nonzero c, $\{x \mid cf(x) < r\} = \{x \mid f(x) < r/c\}$, then $f + c$ and cf are measurable if and only if f is measurable. *What if $c = 0$?*

If $f(x) + g(x) < r$, then $f(x) < r - g(x)$. Therefore, there is rational number p so that $f(x) < p < r - g(x)$. *Why?* It follows that

$$\{x \mid f(x) + g(x) < r\} = \bigcup_{p \in \mathbb{Q}} (\{x \mid f(x) < p\} \cap \{ x \mid g(x) < r - p\})$$

which is a countable union of measurable sets, and hence measurable. Thus $f + g$ is measurable. The measurability of $f - g$ follows similarly.

If $r \geq 0$, then

$$\{x \mid f^2(x) > r\} = \{x \mid f(x) > \sqrt{r}\} \cup \{ x \mid f(x) < -\sqrt{r}\}$$

and if $r < 0$, then

$$\{x \mid f^2(x) > r\} = \operatorname{dom}(f)$$

Hence, f^2 is measurable when f is.

The identity

$$fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right)$$

shows that fg is measurable when f and g are measurable. ∎

Since integration is our goal, we also need to consider sequences and limits.

Theorem 3.18 *Suppose $\{f_n\}$ is a sequence of measurable functions with a common domain. Then $\sup\{f_1, f_2, \dots f_n\}, \sup_n f_n, \inf\{f_1, f_2, \dots f_n\}, \inf_n f_n, \limsup_n f_n$, and $\liminf_n f_n$ are all measurable functions.*

Proof: Let $f = \sup\{f_1, f_2, \dots f_n\}$. Then for any $r \in \mathbb{R}$,

$$\{x \mid f(x) > r\} = \bigcup_{k=1}^{n} \{x \mid f_k(x) > r\}$$

Explain why the set equality above is true! Since the finite union of measurable sets is measurable, f is a measurable function.

Let $g = \sup_n f_n$. Similarly, for any $r \in \mathbb{R}$,

$$\{x \mid g(x) > r\} = \bigcup_{k=1}^{\infty} \{x \mid f_k(x) > r\}$$

The countable union of measurable sets being measurable implies g is a measurable function.

Let $h = \limsup_n f_n$. Then h is measurable because

$$\limsup_n f_n = \inf_n \left(\sup_{k \geq n} f_k \right)$$

Proofs of the statements with infima are left to the exercises. ∎

Previously, we noted the characteristic function of $\mathbb{N}$ was zero almost everywhere, abbreviated as $\chi_{\mathbb{N}} = 0$ a.e., as an example of "equality almost everywhere." If two functions are equal almost everywhere and one is measurable, then so is the other.

Theorem 3.19 *Suppose f is a measurable function and $f = g$ a.e. Then g is a measurable function.*

Proof: Let $E = \{x \mid f(x) \neq g(x)\}$. Then for $r \in \mathbb{R}$,

$$\{x \mid g(x) > r\} = \{x \mid f(x) > r\} \cup \{x \in E \mid g(x) > r\} - \{x \in E \mid g(x) \leq r\}$$

Explain why the set equality above holds! Both $\{x \in E \mid g(x) > r\}$ and $\{x \in E \mid g(x) \leq r\}$ are subsets of E, a set of measure zero. Therefore, both sets are measurable. Hence, if f is measurable, then so is g. ∎

Extend this concept to convergence.

Definition 3.6 (Convergence Almost Everywhere) *A sequence* $\{f_n\}$ converges to f almost everywhere *if and only if the set of values for which $\{f_n\}$ does not converge to f has measure zero. We write $f_n \to f$ a.e.*

■ EXAMPLE 3.8

Define $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} n & x \in \mathbb{Q} \\ \dfrac{1}{n} & \text{otherwise} \end{cases}$$

Then $f_n \to 0$ a.e. on $[0,1]$. *Verify this!* ∎

Measurability is a very general property for a function; however, we will see that a measurable function cannot be too unstructured.

Theorem 3.20 *A function $f : [a,b] \to \mathbb{R}$ is measurable if and only if there is a sequence of simple functions $\{\psi_n\}$ converging to f almost everywhere.*

Proof: Suppose f is measurable. Without loss of generality, assume that f is nonnegative. [Otherwise, write $f = f^+ - f^-$ where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$, and apply the argument below to each part.]

For $n \in \mathbb{N}$, define $A_{n,k}$ by

$$A_{n.k} = \left\{ x \;\middle|\; \frac{k-1}{2^n} \le f(x) < \frac{k}{2^n} \right\}$$

for $k = 1, 2, \ldots, n2^n$, and

$$A_{0,n} = [a,b] - \bigcup_{k=1}^{2^n} A_{n,k}$$

Now let

$$\psi_n(x) = n\chi_{A_{0,n}}(x) + \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{A_{n,k}}(x)$$

Figure 3.5 shows a function and its approximation by ψ_n. It's not difficult to see (*Verify these!*)

1. $\psi_1(x) \le \psi_2(x) \le \cdots$,
2. If $0 \le f(x) \le n$, then $|f(x) - \psi_n(x)| < 2^{-n}$,
3. $\lim_{n\to\infty} \psi_n(x) = f(x)$ a.e.

which proves this direction of the theorem.

Now suppose that a sequence of simple functions $\{\psi_n\}$ converges to f a.e. Since simple functions are measurable, then their limit g is measurable. Since $g = f$ a.e., then f is measurable. ■

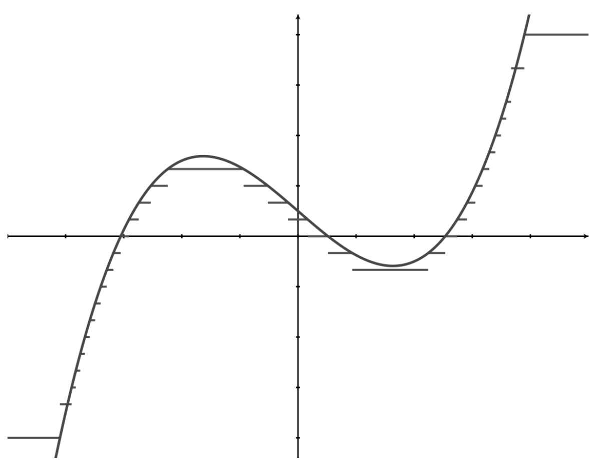


Figure 3.5 A Simple Approximation

We have actually shown that a nonnegative measurable function is the limit of a monotonically increasing sequence of simple functions. It can also be shown that a function is measurable if and only if it is the limit of step functions almost everywhere. Luzin's theorem (1912) showing measurable functions are nearly continuous is a direct consequence of the characterization of measurable functions in terms of step functions.

Theorem 3.21 (Luzin's Theorem) *Let f be a measurable function on $[a, b]$. For any $\epsilon > 0$, there is a continuous function g on $[a, b]$ such that*

$$\mu\left(\{x \mid f(x) \neq g(x)\}\right) < \epsilon$$

We'll defer proving Luzin's theorem as it requires the step function characterization and a theorem of Egorov that we'll see later.

We close our discussion of measurable functions with an interesting observation. A continuous function of a measurable function is measurable. However, a measurable function of a measurable function is not necessarily measurable. In fact, a measurable function of a continuous function need not be measurable.

Now that we've studied measurable functions, it's time to consider integration.

The Lebesgue Integral

Lebesgue's integral is based on partitioning the range, rather than the domain. In his 1926 address to the *Société Mathématique* in Copenhagen, Lebesgue said,

> Let us proceed according to the goal to be attained: to gather or group values of $f(x)$ which differ by little. It is clear then, that we must partition not (a, b), but rather the interval $(\underline{f}, \overline{f})$ bounded by the lower and upper bounds of f on (a, b).

[See Chae (1995, p. 234–248) for an English translation of Lebesgue's address.] Lebesgue used the diagram in Figure 3.6 to illustrate his idea. The range of f is partitioned. The inverse image of the y-axis interval consists of the four components appearing on the x-axis. If L is the length of the inverse image set, then the integral will be bounded by $\underline{f}_n \times L$ and $\overline{f}_n \times L$ where $\underline{f}_n$ and $\overline{f}_n$ are bounds of f on (y_n, y_n). Needing the length L led Lebesgue to develop measure.

Let's build an integral.

Integrals of Bounded Functions There are properties of the Riemann integral that are very useful that we wish to keep when extending the definition. The Riemann integral of a characteristic function of an interval is the interval's length; i.e., $\int_a^b dx = b - a$. The counterpart of an interval is a measurable set. Also, the Riemann integral is linear: $\int_b^a (rf(x) + sg(x))\,dx = r\int_b^a f(x)\,dx + s\int_a^b g(x)\,dx$. We build in these properties by basing the new integral on simple functions. To avoid dealing with infinities, we'll require the measure of the set $\{x \mid \psi(x) \neq 0\}$ to be finite; equivalently, ψ must be zero outside a set of finite measure. (Saying a set has finite measure implicitly assumes the set is measurable.) Recall that simple functions are measurable by definition.

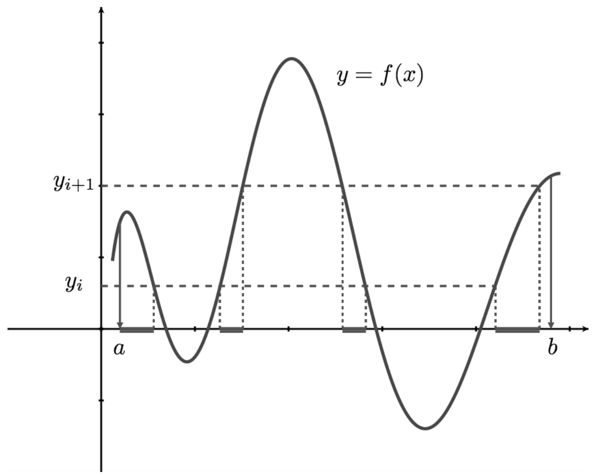


Figure 3.6 Lebesgue's Diagram

Definition 3.7 (Lebesgue Integral of a Bounded Function) *Let ψ be a simple function that is zero outside a set of finite measure, and let $\psi(x) = \sum_{k=1}^{n} a_k \chi_{A_k}(x)$ where each $A_k = \psi^{-1}(a_k)$ is the canonical representation of ψ. Define the* Lebesgue integral of ψ *by*

$$\int \psi \, d\mu = \sum_{k=1}^{n} a_k \mu(A_k)$$

If E is a measurable set, define the Lebesgue integral of ψ over E *by*

$$\int_E \psi \, d\mu = \int \psi \cdot \chi_E \, d\mu = \sum_{k=1}^{n} a_k \mu(A_k \cap E)$$

When the context is clear, we abbreviate the integrals as $\int \psi$ or $\int_E \psi$. Other notations that appear in the literature include $\int \psi(x)\, d\mu(x)$ and $\int \psi(x)\, d\mu$.

■ **EXAMPLE 3.9**

Dirichlet's "monster," the characteristic function of the rationals $\chi_{\mathbb{Q}}$ was the first non-Riemann integrable function on $[0, 1]$ we saw. Now, however, we have

$$\int_{[0,1]} \chi_{\mathbb{Q}} \, d\mu = \mu\left(\mathbb{Q} \cap [0, 1]\right) = 0$$

This function, even though discontinuous everywhere, is Lebesgue integrable! Further, if E is any measurable set, then $\mu(\mathbb{Q} \cap E) = 0$. Therefore we also have $\int_E \chi_{\mathbb{Q}} \, d\mu = 0$. ■

Any simple function has multiple representations, so we need to demonstrate that all representations lead to the same value for the integral.

Theorem 3.22 *Let $\psi(x) = \sum_{k=1}^{n} a_k \chi_{E_k}(x)$ where each E_k has finite measure and the E_k are pairwise disjoint; that is, $E_k \cap E_j = \emptyset$ for $k \neq j$. Then*

$$\int \psi \, d\mu = \sum_{k=1}^{n} a_k \mu(E_k)$$

Proof: For any r, the set $\psi^{-1}(r) = \bigcup_{a_k = r} E_k$. Since Lebesgue measure is additive for disjoint sets, $\mu\left(\psi^{-1}(r)\right) = \sum_{a_k=r} \mu(E_k)$. Set $R = \operatorname{range}(\psi)$. Then R is finite. *Why?* Therefore

$$\begin{aligned}\int \psi \, d\mu &= \sum_{r \in R} r \mu\left(\psi^{-1}(r)\right) \\ &= \sum_{r \in R} r \cdot \left(\sum_{a_k = r} \mu(E_k) \right) \\ &= \sum_{k=1}^{n} a_k \mu(E_k)\end{aligned}$$

Thus the integral of ψ is independent of representation. ■

A development of the Lebesgue integral using partitions closely following the Riemann integral is given in Bear (1995).

Lebesgue integrals are linear and monotone by design.

Theorem 3.23 *Let ψ and ϕ be two simple functions that are both zero outside a set of finite measure, and let a and b be real constants. Then*

$$\int (a\psi + b\phi) d\mu = a \int \psi \, d\mu + b \int \phi \, d\mu$$

If $\psi(x) \leq \phi(x)$ a.e., then

$$\int \psi \, d\mu \leq \int \phi \, d\mu$$

Proof: We need to start by dealing with the simple function $a\psi + b\phi$. Let the canonical representation of ψ be $\psi(x) = \sum_{k=1}^{N_1} a_k \chi_{A_k}(x)$ and that of ϕ be $\phi(x) = \sum_{k=1}^{N_2} b_k \chi_{B_k}(x)$. Set $A_0 = \psi^{-1}(0)$ and $B_0 = \phi^{-1}(0)$. Form the set

$$\{E_k \,|\, k = 1, \ldots, N\} = \{A_i \cap B_j \,|\, i = 0, \ldots, N_1, j = 0, \ldots, N_2\}$$

Since $\{A_k\}$ and $\{B_k\}$ are pairwise disjoint sets, then $\{E_k\}$ is a pairwise disjoint collection. Thus,

$$\psi(x) = \sum_{k=1}^{N} a_k \chi_{E_k}(x) \qquad \text{and} \qquad \phi(x) = \sum_{k=1}^{N} b_k \chi_{E_k}(x)$$

Then we see that

$$a\psi(x) + b\phi(x) = \sum_{k=1}^{N} (aa_k + bb_k)\chi_{E_k}(x)$$

It now follows from the previous theorem that $\int (a\psi + b\phi) d\mu = a \int \psi\, d\mu + b \int \phi\, d\mu$.

Now, if $\psi(x) \leq \phi(x)$ a.e., then $\int \phi\, d\mu - \int \psi\, d\mu = \int (\phi - \psi)\, d\mu$ by part 1. Since $\phi(x) - \psi(x) \geq 0$ a.e., the definition of the integral gives $\int (\phi - \psi)\, d\mu \geq 0$, and part 2 is seen to hold. ■

Since the Lebesgue integral is linear, we don't need the requirement that the sets be disjoint in the representation of a simple function. *Why?*

In defining the Riemann integral, we used the maximum and minimum of the function f on the subinterval $[x_{k-1}, x_k]$ to build Riemann-Darboux sums. The upper Riemann integral is

$$\overline{\int_a^b} f(x)\, dx = \inf_P \sum_{k=1}^{n} \max_{[x_{k-1}, x_k]} f(x)\, \Delta x_k$$

We can look at this sum as the integral of a step function. *How?* Then

$$\overline{\int_a^b} f(x)\, dx = \inf_{\phi \geq f} \int_a^b \phi(x)\, dx$$

where ϕ is a step function. We use this reasoning, extend from step functions to simple functions, to build the Lebesgue integral of a function. For a bounded real-valued function f on a measurable set E, we check whether, for simple functions ψ and ϕ, the two expressions

$$\inf_{\psi \geq f} \int_E \psi\, d\mu \qquad \text{and} \qquad \sup_{\phi \leq f} \int_E \phi\, d\mu$$

are equal.

Theorem 3.24 *Suppose F is a bounded real-valued function on E, a set with finite measure. Then for all simple functions ψ and ϕ,*

$$\inf_{\psi \geq f} \int_E \psi\, d\mu = \sup_{\phi \leq f} \int_E \phi\, d\mu$$

if and only if f is measurable.

Proof: First, suppose that f is bounded by M and measurable. For $k = -n, -n+1, -n+2, \ldots, n-1, n$, put

$$E_k = \left\{ x \;\middle|\; \frac{k-1}{n} \cdot M < f(x) \leq \frac{k}{n} \cdot M \right\}$$

The E_k are measurable (*Why?*), disjoint, and $E = \bigcup_k E_k$. Therefore

$$\sum_{k=-n}^{n} \mu(E_k) = \mu(E)$$

Set

$$\psi_n(x) = \sum_{k=-n}^{n} \frac{kM}{n} \chi_{E_k}(x) \quad \text{and} \quad \phi_n(x) = \sum_{k=-n}^{n} \frac{(k-1)M}{n} \chi_{E_k}(x)$$

The definition of ψ and ϕ implies that

$$\phi_n(x) \le f(x) \le \psi_n(x)$$

and $\psi_n(x) - \phi_n(x) \ge 0$. *Verify this!* Thence

$$\inf \int_E \psi \, d\mu \le \int_E \psi_n \, d\mu = \sum_{k=-n}^{n} \frac{kM}{n} \mu(E_k)$$

and

$$\sup \int_E \phi \, d\mu \ge \int_E \phi_n \, d\mu = \sum_{k=-n}^{n} \frac{(k-1)M}{n} \mu(E_k)$$

Subtracting the second inequality from the first and remembering $\phi_n - \psi_n \ge 0$ for all n, we arrive at

$$0 \le \inf \int_E \psi \, d\mu - \sup \int_E \phi \, d\mu \le \frac{M}{n} \sum_{k=-n}^{n} \mu(E_k) = \frac{M}{n} \mu(E)$$

Since n is arbitrary, this relation holds for all n. Hence

$$\inf \int_E \psi \, d\mu = \sup \int_E \phi \, d\mu$$

Now, for the other direction, we assume that $\inf \int_E \psi \, d\mu = \sup \int_E \phi \, d\mu$. For any n, there are simple functions ψ_n and ϕ_n such that $\phi_n(x) \le f(x) \le \psi_n(x)$ and

$$\int \psi_n \, d\mu - \int \phi_n \, d\mu < \frac{1}{n}$$

Why? The functions

$$\hat{\psi} = \inf \psi_n \qquad \text{and} \qquad \hat{\phi} = \sup \phi_n$$

are measurable and $\hat{\phi}(x) \le f(x) \le \hat{\psi}(x)$. Let $F = \{x \mid \hat{\phi}(x) < \hat{\psi}(x)\}$ and $F_r = \{x \mid \hat{\phi}(x) < \hat{\psi}(x) - 1/r\}$. It follows that $F = \bigcup_r F_r$. Since

$$\phi_n(x) \le \hat{\phi}(x) \le f(x) \le \hat{\psi}(x) \le \psi_n(x)$$

we have $F_r \subseteq \{x \mid \phi_n(x) < \psi_n(x) - 1/r\}$ which has measure less than r/n. *Why?* Then $\mu(F_r) = 0$ as n is arbitrary. Hence $\mu(F) = 0$, and $\hat{\phi} = \hat{\psi}$ a.e. Therefore $\hat{\phi} = f$ a.e., which implies that f is measurable. ∎

So, if a function f is measurable and bounded, then the analogues of the upper and lower integrals are equal. We can use this fact to define the integral of f.

Definition 3.8 *If f is a bounded measurable function defined on a set E having finite measure, then the* Lebesgue integral of f over E *is*

$$\int_E f \, d\mu = \inf \int_E \psi \, d\mu$$

for all simple functions $\psi \geq f$.

We already know that Lebesgue's integral can handle functions that Riemann's cannot. Is every Riemann integrable function also Lebesgue integrable? Recall we interpreted Riemann sums as integrals of step functions. The answer now comes easily.

Theorem 3.25 *Let $f : [a, b] \to \mathbb{R}$ be bounded and Riemann integrable on $[a, b]$. Then f is measurable and*

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

Proof: Every step function is a simple function. Interpret Riemann sums as step functions to yield

$$\underline{\int_a^b} f(x) \, dx \leq \sup_{\phi \leq f} \int_{[a,b]} \phi \, d\mu \leq \inf_{\psi \geq f} \int_{[a,b]} \psi \, d\mu \leq \overline{\int_a^b} f(x) \, dx$$

The inequalities are actually equalities given the Riemann integrability of f. Then f is measurable by Theorem 3.24. ∎

This result tells us that Lebesgue's integration is an extension of Riemann's: we have added to the class of integrable functions. Since the Lebesgue and Riemann integrals are equal, we normally write $\int_a^b f \, d\mu$ rather than $\int_{[a,b]} f \, d\mu$.

Let's collect several properties that follow immediately from the definition.

Theorem 3.26 (Properties of the Lebesgue Integrals of Bounded Functions) *Let f and g be bounded measurable functions both defined on a set E of finite measure, and let $c \in \mathbb{R}$. Then*

1. $\displaystyle\int_E cf \, d\mu = c \int_E f \, d\mu.$
2. $\displaystyle\int_E f + g \, d\mu = \int_E f \, d\mu + \int_E g \, d\mu.$

3. *If* $f \leq g$ *a.e., then* $\int_E f\,d\mu \leq \int_E g\,d\mu$.
4. *If* $f = g$ *a.e., then* $\int_E f\,d\mu = \int_E g\,d\mu$.
5. *If* $m \leq f(x) \leq M$*, then*

$$m \cdot \mu(E) \leq \int_E f\,d\mu \leq M \cdot \mu(E)$$

6. *If* E_1 *and* E_2 *are disjoint measurable sets with* $E_1 \cup E_2 = E$*, then*

$$\int_{E_1} f\,d\mu + \int_{E_2} f\,d\mu = \int_E f\,d\mu$$

The proofs are straightforward and are in Exercise 3.22.

Corollary 3.27 *If* f *is a bounded measurable function defined on a set* E *of finite measure, then*

$$\left|\int_E f\,d\mu\right| \leq \int_E |f|\,\,d\mu$$

We end this segment with a theorem that characterizes Riemann integration by using Lebesgue's theory. Riemann's condition for integrability is based on the sets on which a function varies more than a small amount. If the total length of these sets is small enough, the function is integrable. The *oscillation of a function* f *on the interval* I is given by

$$\omega(f, I) = \sup_{s,t \in I} |f(s) - f(t)|$$

and the *oscillation of* f *at the point* x is

$$\omega(f, x) = \inf_{\delta > 0} \omega(f, N_\delta(x))$$

Recall $N_\delta(x) = (x - \delta, x + \delta)$. Riemann's theorem is stated here without proof for comparison to Lebesgue's condition.

Theorem 3.28 (Riemann's Criterion for Riemann Integrability) *Let the function* f *be bounded on the interval* $[a, b]$*. Then* f *is Riemann integrable if and only if for any* $\epsilon, \sigma > 0$ *there is a* $\delta > 0$ *such that for any partition* P *with* $\|P\| < \delta$ *it follows that*

$$\sum_{\{j \,|\, \omega(f, I_j) > \sigma\}} \Delta x_j < \epsilon$$

where $I_j = [x_{j-1}, x_j]$.

Riemann's criterion in words is: f is integrable if and only if the sum of the lengths of the intervals where f varies more than ϵ is less than σ whenever the partition's norm is less than δ. That is, with tongue firmly in cheek, f is integrable if and only if it misbehaves in only a small collection of tiny neighborhoods.

Lebesgue recognized that Riemann's condition was really characterizing continuity. A function needed to be "almost continuous" to be Riemann integrable.

Theorem 3.29 (Lebesgue's Criterion for Riemann Integrability) *Let the function f be bounded on the interval $[a, b]$. Then f is Riemann integrable if and only if f is continuous almost everywhere on $[a, b]$.*

Proof: First, suppose f is Riemann integrable and let $\epsilon > 0$. Define the set N_r by

$$N_r = \{x \mid \omega(f, x) > r\}$$

Then any interval I that contains a point of N_r will have $\omega(f, I) \geq r$. Since f is Riemann integrable, there is a partition P so that the difference between the upper and lower Riemann sums is less than $r\epsilon/2$. Set $I_j = [x_{j-1}, x_j]$. Then

$$\sum_{\{j \mid I_j \cap N_r \neq \emptyset\}} \omega(f, I_j)\Delta x_j \leq \sum_{k=1}^{n} \omega(f, I_j)\Delta x_j < \frac{r\epsilon}{2}$$

Now, since the sum is less than $r\epsilon/2$, and $\omega(f, I_j) > r$ (*Why?*), we must have

$$\sum_{\{j \mid I_j \cap N_r \neq \emptyset\}} \Delta x_j < \frac{r\epsilon}{2r} = \frac{\epsilon}{2}$$

Each point of N_r is in some $I_j = [x_{j-1}, x_j]$; we can cover N_r with the collection of open sets $\{(x_{j-1}, x_j) \mid (x_{j-1}, x_j) \cap N_r \neq \emptyset\} \cup \{(x_j - \epsilon/(2n), x_j + \epsilon/(2n))\}$ which has a total length less than ϵ. Since $\epsilon > 0$ is arbitrary, N_r must have measure zero. The set N of points of discontinuity of f is the countable union $\bigcup_{k=1}^{\infty} N_{1/k}$. Therefore N has measure zero.

For the other direction, suppose f is continuous almost everywhere on $[a, b]$ and $|f(x)| \leq M$. Let $\epsilon > 0$ be given. This part of the proof uses Riemann's idea of breaking the sum into two parts, one with small oscillations, the other with small intervals. Set $r = \epsilon/(2(b - a))$, and now let

$$N_r = \{x \mid \omega(f, x) \geq r\}$$

Thus N_r is a closed set of measure zero. *Why?* Hence we can cover N_r with a finite set $\mathcal{F}$ of disjoint closed intervals of total length less than $\epsilon/(4M + 1)$. Any point x of $[a, b]$ not covered by these intervals has oscillation less than $r = \epsilon/(2(b - a))$. Choose x_k from the remaining points to add to the endpoints of the intervals in $\mathcal{F}$ to form a partition of $[a, b]$. Intervals of the partition are either in $\mathcal{F}$ or have

$(\sup f - \inf f) \le \epsilon/(2(b-a))$. Then

$$\begin{aligned}\sum_{k=1}^{n}\left(\sup_{I_k} f(x) - \inf_{I_k} f(x)\right)\Delta x_k &\le \sum_{\{k \mid I_k \notin \mathcal{F}\}} \frac{\epsilon}{2(b-a)}\,\Delta x_k + \sum_{\{k \mid I_k \in \mathcal{F}\}} 2M\Delta x_k \\ &\le \frac{\epsilon}{2(b-a)} \sum_{k=1}^{n} \Delta x_k + 2M \sum_{\{k \mid I_k \in \mathcal{F}\}} \Delta x_k \\ &< \frac{\epsilon}{2} + 2M \cdot \frac{\epsilon}{4M+1} < \epsilon\end{aligned}$$

Then, by Theorem 2.36, f is Riemann integrable. ■

With the results for bounded functions in hand, we can extend integration to unbounded functions, the analogue of improper Riemann integrals.

Integrals of Unbounded Functions First, we define Lebesgue integrals of nonnegative measurable functions that may be unbounded; then we deal with general measurable functions by splitting the function into positive and negative parts, working with each separately. A function may be unbounded in different ways: a function may have an unbounded range or an unbounded domain. We focus on functions with an unbounded range; analogous techniques may be applied to handle unbounded domains.

Define the n-*truncation* of a nonnegative function f on the interval $[a, b]$ to be the minimum of $f(x)$ and n. Clearly, the sequence $\{f_n\}$ of n-truncations converges monotonically to f on $[a, b]$. Now $f_n = \min\{f, n\}$ is measurable and bounded for each n. *Why?* Hence $\int_a^b f_n \, d\mu$ exists for all n.

Definition 3.9 (Lebesgue Integral of an Unbounded Function) *Let f be a nonnegative measurable function on $[a, b]$, and let $\{f_n\}$ be the sequence of n-truncations of f. Then*

$$\int_a^b f \, d\mu = \lim_{n\to\infty} \int_a^b f_n \, d\mu$$

which may be infinite. If the limit is finite, we say f is Lebesgue integrable *on $[a, b]$.*

■ EXAMPLE 3.10

Determine $\displaystyle\int_0^2 f \, d\mu$ for

$$f(x) = \begin{cases} \dfrac{1}{(x-1)^{2/3}} & x \ne 1 \\ +\infty & x = 1 \end{cases}$$

The n-truncations of f are

$$f_n(x) = \begin{cases} \dfrac{1}{(x-1)^2} & 0 \le x < 1 - \dfrac{1}{n^{3/2}} \\ n & 1 - \dfrac{1}{n^{3/2}} \le x \le 1 + \dfrac{1}{n^{3/2}} \\ \dfrac{1}{(x-1)^2} & 1 + \dfrac{1}{n^{3/2}} < x \le 2 \end{cases}$$

Graph f and several f_n. Then f_n is continuous, and hence Riemann integrable. Thus

$$\begin{aligned} \int_0^2 f_n \, d\mu &= \int_0^{1-1/n^{3/2}} f_n \, d\mu + \int_{1-1/n^{3/2}}^{1+1/n^{3/2}} n \, d\mu + \int_{1+1/\sqrt{n^3}}^{2} f_n \, d\mu \\ &= \left(3 - \frac{3}{\sqrt{n}}\right) + \frac{2}{\sqrt{n}} + \left(3 - \frac{3}{\sqrt{n}}\right) \\ &= 6 - \frac{4}{\sqrt{n}} \end{aligned}$$

Letting n go to infinity gives us $\displaystyle\int_0^2 f \, d\mu = 6$. ■

Does the improper Riemann integral of f from Example 3.10 *converge?*

Does the algebra of integrals still apply to unbounded nonnegative functions? Yes!

Theorem 3.30 *If f and g are nonnegative measurable functions, and $c > 0$, then*

1. $\displaystyle\int_a^b cf \, d\mu = c \int_a^b f \, d\mu,$

2. $\displaystyle\int_a^b f + g \, d\mu = \int_a^b f \, d\mu + \int_a^b g \, d\mu.$

3. *If $f \le g$ a.e., then* $\displaystyle\int_a^b f \, d\mu \le \int_a^b g \, d\mu.$

The proofs are straightforward applications of Theorem 3.26 and are left to the exercises.

For functions that are both positive and negative we use the device from the proof of Theorem 3.20 of splitting the function into positive and negative parts. Define the *positive part of f* to be $f^+(x) = \max\{f(x), 0\}$ and the *negative part of f* to be $f^-(x) = \max\{-f(x), 0\}$. Both f^+ and f^- will be nonnegative functions.

Definition 3.10 *Let f be a measurable function on $[a,b]$. Then*

$$\int_a^b f\,d\mu = \int_a^b f^+\,d\mu - \int_a^b f^-\,d\mu$$

which may be infinite. If the integrals of f^+ and f^- are both finite, so is the integral of f. Then we say f is Lebesgue integrable *on $[a,b]$, and write $f \in \mathcal{L}([a,b])$. If both $\int_a^b f^+\,d\mu$ and $\int_a^b f^-\,d\mu$ are infinite, we say the integral of f does not exist.*

Combine the two facts

$$f \in \mathcal{L}([a,b]) \iff f^+ \text{ and } f^- \in \mathcal{L}([a,b])$$
$$|f| = f^+ + f^-$$

to see that f is Lebesgue integrable if and only if $|f|$ is. Is this relation also true for Riemann inegrals? Check what happens with the function $R(x) = 1$ if $x \in \mathbb{Q}$ and $R(x) = -1$ otherwise.

The standard way to integrate f over a measurable subset of $[a,b]$ is to multiply f by the characteristic function of the set.

Definition 3.11 *If f is a measurable function and E is a measurable subset of $[a,b]$, then*

$$\int_E f\,d\mu = \int_a^b f \cdot \chi_E\,d\mu$$

Observe that a set of measure zero will not affect the value of a Lebesgue integral since

$$\mu(E) = 0 \implies \int_E f\,d\mu = 0$$

Verify this! Dirichlet's and Thomae's functions both have Lebesgue integral equal to zero over $[0,1]$ (or over any finite interval for that matter). Compare this result to Riemann integrals of each: Dirichlet's function is *not* Riemann integrable, while Thomae's is! *Why doesn't this contradict Theorem* 3.29*?*

The expected theorem on the algebra of general Lebesgue integrals is left to the reader as an exercise.

We turn to studying the advantages for convergence gained by Lebesgue's integral.

3.3 MEASURE, INTEGRAL, AND CONVERGENCE

One of the main motivations in the development of integration was the question of convergence of Fourier series. The expression

$$S_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,\frac{\sin\big((n+\frac{1}{2})(x-t)\big)}{\sin\big(\frac{1}{2}(x-t)\big)}\,dt$$

is Dirichlet's integral form of the nth partial sum of f's Fourier series. Letting f_n be the integrand brings us to question what properties does f need to have to guarantee convergence for S_n. We are led to investigating the relation between integration and sequences of functions.

To learn more about how Fourier series influenced the development of real analysis, see Bressoud (2005, 2008). A very readable exposition of Fourier series appears in Jackson (1941); also look ahead at Section 4.3 for an introduction.

Types of Convergence

There are several different modes of convergence that we are led to consider. In beginning real analysis, we focused on pointwise and uniform convergence. We have seen convergence almost everywhere earlier in this chapter. Now we define new types by changing how we measure the "distance" between functions.

Definition 3.12 (Types of Convergence) *Let $\{f_n\}$ be a sequence of functions. Then*

Almost Everywhere: *f_n converges to f* almost everywhere *if and only if $f_n(x) \to f(x)$ except on a set of measure zero.*

Almost Uniformly: *f_n converges to f* almost uniformly *if and only if for every $\epsilon > 0$ there is a set E_ϵ of measure less than ϵ such that $f_n \to f$ uniformly on E_ϵ^c, the complement of E_ϵ.*

In Mean: *f_n converges to f* in mean *if and only if*

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0$$

In Measure: *f_n converges to f* in measure *if and only if for every $\epsilon > 0$*

$$\lim_{n\to\infty} \mu \{x \mid |f_n(x) - f(x)| > \epsilon\} = 0$$

Figure 3.7 summarizes the relation between the different modes of convergence. Each arrow in the diagram represents a convergence theorem. For instance, almost uniform convergence implies both convergence almost everywhere and convergence in measure. If we restrict sequences to bounded domains like $[a, b]$, several more arrows could be drawn. Soon we'll add an arrow with a 1911 theorem of Egorov's. See Bressoud (2008) or Royden (1988) for more details and further theory.

■ **EXAMPLE 3.11**

1. Define the sequence $\{f_n\}$ by $f_1 = \chi_{[0,1]}$, $f_2 = \chi_{[0,1/2]}$, $f_3 = \chi_{[1/2,1]}$, $f_4 = \chi_{[0,1/3]}$, $f_5 = \chi_{[1/3,2/3]}$, $f_6 = \chi_{[2/3,1]}$, and so forth. Figure 3.8 shows graphs of f_1 through f_6. Then, following the order diagrammed in Figure 3.7,

   (a) f_n does not converge uniformly on $[0, 1]$.

   (b) $f_n \to 0$ in mean on $[0, 1]$.

   (c) f_n does not converge almost uniformly on $[0, 1]$.

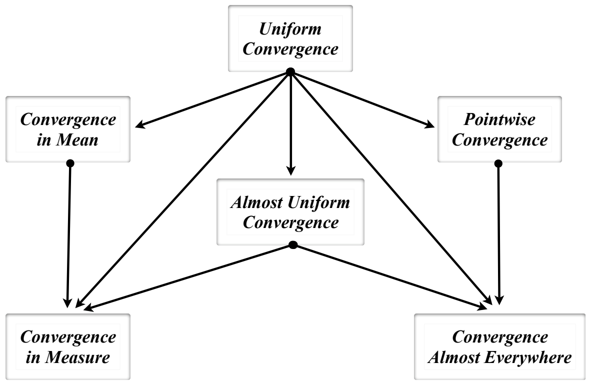


Figure 3.7 Types of Convergence

(d) f_n does not converge pointwise on $[0, 1]$ or even at any $x \in [0, 1]$.

(e) $f_n \to 0$ in measure on $[0, 1]$.

(f) f_n does not converge almost everywhere on $[0, 1]$.

Verify each claim above!

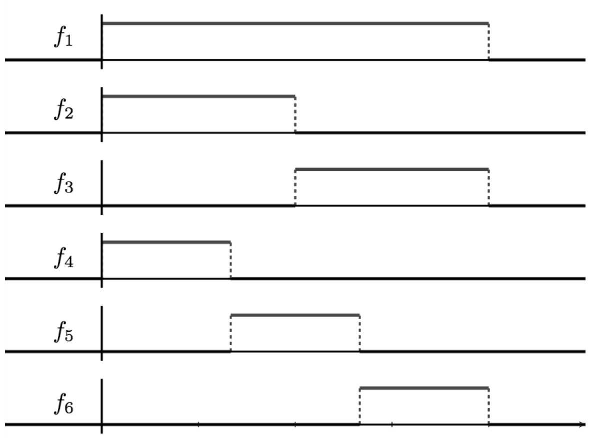


Figure 3.8 Convergence in Mean and Measure

2. Let

$$g_n(x) = \frac{n^2 x}{1 + n^3 x^2}$$

Graphs of $g_n(x)$ are in Figure 3.9. Then

(a) g_n does not converge uniformly on $[0, 1]$.

(b) $g_n \to 0$ in mean on $[0, 1]$.

(c) $g_n \to 0$ almost uniformly on $[0, 1]$. *Show that* $\{g_n\}$ *converges uniformly on* $[\epsilon, 1] = [0, 1] - N_\epsilon(0)$ *for any* $\epsilon > 0$.

(d) $g_n \to 0$ pointwise on $[0, 1]$.

(e) $g_n \to 0$ in measure on $[0, 1]$.

(f) $g_n \to 0$ almost everywhere on $[0, 1]$.

Verify each claim above!

■

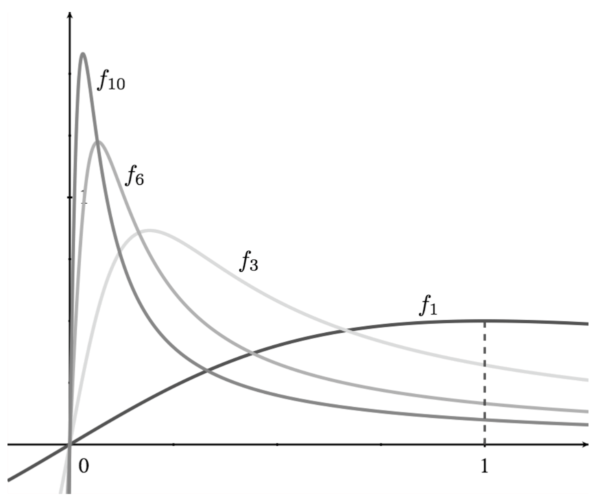


Figure 3.9 Almost Uniform, But Not Uniform Convergence

There is a general discussion of different types of convergence and their relationships in Bartle (1995, Chapter 7).

The first relation we would add to the diagram is: converging pointwise is nearly converging uniformly for a sequence of functions defined on a bounded interval $[a, b]$.

Theorem 3.31 (Egorov's Theorem) *If the sequence of measurable functions* $\{f_n\}$ *converges to* f *almost everywhere on* $[a, b]$, *then it converges almost uniformly to* f *on* $[a, b]$; *that is, for any* $\epsilon > 0$, *there is a set* $E \subset [a, b]$ *with* $\mu(E) < \epsilon$ *such that* f_n *converges to* f *uniformly on* $[a, b] - E$.

Proof: Following Bressoud (2008, p. 193), we do two reductions to simplify the proof. First, $f_n \to f$ a.e. on $[a, b]$ if and only if $f_n - f \to 0$ a.e. on $[a, b]$. Thus, we can assume $f = 0$ without loss of generality. Second, define $g_n(x) = \sup_{m \geq n} |f_m(x)|$. Then $g_n \to 0$ uniformly on E if and only if $f_n \to 0$ uniformly on E. *Verify this!* Since $\{g_n\}$ is made of supremums of smaller and smaller sets, they must monotonically decrease to zero. Let A be the set of measure zero where $f_n \not\to 0$; therefore $g_n \not\to 0$ on A.

Define

$$E_{k,n} = \left\{ x \,|\, 0 \leq g_n(x) < \frac{1}{2^k} \right\}$$

Then each $E_{k,n}$ is measurable, and

$$E_{k,1} \subseteq E_{k,2} \subseteq E_{k,3} \subseteq \cdots$$

because g_n monotonically decreases to zero. Further, since $g_n \to 0$ on $[a, b] - A$, we have that

$$[a, b] - A \subseteq \bigcup_{n=1}^{\infty} E_{k,n}$$

Since $\mu(A) = 0$, and the limit of the measures of nested sets is the measure of the limit of the sets (*By which theorem?*),

$$b - a \leq \mu\left(\bigcup_{n=1}^{\infty} E_{k,n} \right) = \lim_{n \to \infty} \mu(E_{k,n})$$

Let $\epsilon > 0$ be given. For each $k \in \mathbb{N}$, there is an $n = n_k$ such that

$$\mu(E_{k,n}) > (b - a) - \frac{\epsilon}{2^k}$$

Why? Now define

$$E = [a, b] - \bigcap_{k=1}^{\infty} E_{k,n_k} = \bigcup_{k=1}^{\infty} ([a, b] - E_{k,n_k})$$

Then $\mu(E)$ is bounded by

$$\mu(E) \leq \mu\left(\bigcup_{k=1}^{\infty} ([a, b] - E_{k,n_k}) \right) \leq \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon$$

i.e., the measure of E is less than ϵ. For any $x \in [a, b] - E$, we have $x \in \bigcap_{k=1}^{\infty} E_{k,n_k}$. Hence $x \in E_{k,n_k}$ for every k. Therefore, for every $m > n_k$,

$$0 \leq g_m(x) \leq g_{n_k}(x) < \frac{1}{2^k}$$

Thus g_n converges uniformly to zero on $[a, b] - S$, and the theorem is proved. ■

Lebesgue Integration and Convergence

Lebesgue's integral expands the class of functions that are integrable. Does adding more functions to the class change the requirements for interchanging integrals and limits? We'll see the answer gives an enhanced theory of convergence.

If each function in a sequence is bounded by the same value, we call the sequence *uniformly bounded*. Our first convergence theorem is for uniformly bounded sequences.

Theorem 3.32 (Bounded Convergence Theorem) *If $\{f_n\}$ is a uniformly bounded sequence of measurable functions converging to f a.e. on $[a,b]$, then f is measurable and*

$$\lim_{n\to\infty}\int_a^b f_n\,d\mu = \int_a^b \lim_{n\to\infty} f_n\,d\mu = \int_a^b f\,d\mu$$

Proof: On the set of measure zero where $f_n \not\to f$, define f to be zero. Let M be a bound for f_n on $[a,b]$. Since f is the limit of bounded measurable functions, it must be bounded and measurable. Therefore, f is integrable.

Let $\epsilon > 0$. Egorov's theorem implies there is a set E with $\mu(E) < \epsilon$ so that $f_n \to f$ uniformly on $[a-b] - E$. Then, for n large enough,

$$\begin{aligned}
\left|\int_a^b f_n\,d\mu - \int_a^b f\,d\mu\right| &\le \int_a^b |f_n - f|\;d\mu \\
&= \int_E |f_n - f|\;d\mu + \int_{[a,b]-E} |f_n - f|\;d\mu \\
&\le 2M\mu(E) + \epsilon\,\mu\left([a,b]-E\right) \\
&= \left[2M + \mu\left([a,b]-E\right)\right]\epsilon < K\epsilon
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the integrals must be equal. ■

If instead of bounding the functions by a constant, we ask that the functions be nonnegative and bound each by the next in the sequence, we can obtain the same result. This requirement is simply that the sequence is monotone increasing and nonnegative. Beppo Levi proved this theorem in 1906. First, we prove a lemma due to Fatou (1906) that is very useful in itself.

Lemma 3.33 (Fatou's Lemma) *If $\{f_n\}$ is a sequence of nonnegative measurable functions converging to f a.e. on $[a,b]$, then*

$$\int_a^b f\,d\mu = \int_a^b \lim_{n\to\infty} f_n\,d\mu \le \liminf_{n\to\infty}\int_a^b f_n\,d\mu$$

Proof: Without loss of generality, assume $f_n \to f$ for all x in $[a,b]$ as sets of measure zero do not affect integrals. Let ψ be a measurable function on $[a,b]$ that is less than or equal to f and is bounded by, say, M. Set

$$\psi_n(x) = \min\{f_n(x), \psi(x)\} \le f_n(x)$$

Then $\{\psi_n\}$ is a uniformly bounded sequence of measurable functions. By the bounded convergence theorem,

$$\int_a^b \lim_{n\to\infty} \psi_n \, d\mu = \lim_{n\to\infty} \int_a^b \psi_n \, d\mu \le \liminf_{n\to\infty} \int_a^b f_n \, d\mu$$

Take the supremum over all bounded measurable functions ψ such that $\psi \le f$ to obtain

$$\int_a^b f \, d\mu \le \liminf_{n\to\infty} \int_a^b f_n \, d\mu$$

■

Levi's theorem is also called the *monotone convergence theorem.*

Theorem 3.34 (Levi's Theorem) *If $\{f_n\}$ is a monotone increasing sequence of nonnegative measurable functions converging to f a.e. on $[a, b]$, then f is measurable and*

$$\lim_{n\to\infty} \int_a^b f_n \, d\mu = \int_a^b \lim_{n\to\infty} f_n \, d\mu = \int_a^b f \, d\mu$$

Proof: Fatou's lemma gives

$$\int_a^b f \, d\mu \le \liminf_{n\to\infty} \int_a^b f_n \, d\mu$$

On the other hand, for each n, we have $f_n \le f$. Therefore $\int_a^b f_n \, d\mu \le \int_a^b f \, d\mu$. Take the supremum of both sides to see

$$\limsup_{n\to\infty} \int_a^b f_n \, d\mu \le \int_a^b f \, d\mu$$

Thus

$$\int_a^b f \, d\mu = \lim_{n\to\infty} \int_a^b f_n \, d\mu$$

■

A clever proof of Levi's theorem using double sequences of simple functions rather than Fatou's lemma appears in Bressoud (2008, p. 174) and Burk (2007, p. 122).

■ EXAMPLE 3.12

Find $\displaystyle\int_0^1 x \ln\left(1/x^4\right) \, d\mu$ using

$$f_n(x) = \begin{cases} x \ln\left(\dfrac{1}{x^4}\right) & \dfrac{1}{n} \le x \le 1 \\ 0 & 0 \le x < \dfrac{1}{n} \end{cases}$$

Since f_n is nonnegative and monotone increasing (*Show this! Also, graph several f_n's.*), we can apply Levi's theorem. Now

$$\begin{aligned}\int_0^1 f_n \, d\mu &= \int_{1/n}^1 f_n \, d\mu \\ &= \int_{1/n}^1 x \ln\left(\frac{1}{x^4}\right) \, dx\end{aligned}$$

where the last integral is a Riemann integral. *Why is this change valid?* Evaluating, we see

$$\int_0^1 f_n \, d\mu = 1 - \left(\frac{1}{n^2} + \frac{2\ln(n)}{n^2}\right)$$

which goes to 1 as $n \to \infty$. Hence, by Levi's theorem, the original integral is equal to 1. ■

A nice corollary comes from rephrasing Levi's theorem in terms of series.

Corollary 3.35 *Suppose $\{f_n\}$ is a sequence of nonnegative measurable functions on $[a, b]$. Then*

$$\int_a^b \sum_{k=1}^{\infty} f_n(x) \, d\mu = \sum_{k=1}^{\infty} \int_a^b f_n(x) \, d\mu$$

Riemann integrals required uniform convergence in order to exchange of summation and integration.

Our last result on integration and convergence is the very powerful *Lebesgue's dominated convergence theorem* of 1910. Lebesgue was able to replace "monotone and bounded" with "bounded by an integrable function."

Theorem 3.36 (Lebesgue's Dominated Convergence Theorem) *Let $\{f_n\}$ be a sequence of integrable functions converging to f a.e. on $[a, b]$. If there is an integrable function g on $[a, b]$ such that $|f_n| \leq g$ for all n, then f is integrable on $[a, b]$, and*

$$\int_a^b f \, d\mu = \int_a^b \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_a^b f_n \, d\mu$$

Proof: Since f_n is integrable, it is measurable. Therefore f, the limit of the f_n's, is measurable and integrable. As before, we can adjust f on a set of measure zero without loss of generality; thus, we can assume f is finite and is the limit of f_n on $[a, b]$.

Define two monotone sequences, one increasing and one decreasing, by

$$\underline{f}_n = \inf\{f_n, f_{n+1}, \dots\} \quad \text{and} \quad \overline{f}_n = \sup\{f_n, f_{n+1}, \dots\}$$

Then $-g \leq \underline{f}_n \leq f_n \leq \overline{f}_n \leq g$. So both $\underline{f}_n$ and $\overline{f}_n$ are integrable. Now, both sequences $\{g + \underline{f}_n\}$ and $\{g - \overline{f}_n\}$ are nonnegative and monotone increasing. *Why?*

Therefore both converge to integrable functions, namely $g + \underline{f}_n \to g + f$ and $g - \overline{f}_n \to g - f$.

Apply Levi's theorem to see

$$\int_a^b (g+f)\,d\mu = \lim_{n\to\infty} \int_a^b (g + \underline{f}_n)\,d\mu$$
$$= \int_a^b g\,d\mu + \int_a^b \lim_{n\to\infty} \underline{f}_n\,d\mu$$

Subtract the $\int_a^b g\,d\mu$ from both sides to have

$$\int_a^b f\,d\mu = \int_a^b \lim_{n\to\infty} \underline{f}_n\,d\mu$$

Similarly,

$$\int_a^b f\,d\mu = \int_a^b \lim_{n\to\infty} \overline{f}_n\,d\mu$$

On the other hand, $\underline{f}_n \le f_n \le \overline{f}_n$ implies that

$$\int_a^b \underline{f}_n\,d\mu \le \int_a^b f_n\,d\mu \le \int_a^b \overline{f}_n\,d\mu$$

for all n. Hence $\int_a^b f\,d\mu = \lim_{n\to\infty} \int_a^b f_n\,d\mu$, and the theorem holds. ∎

■ EXAMPLE 3.13

Show that

$$\lim_{n\to\infty} \int_0^1 \frac{nx}{1+n^2x^2}\,d\mu = 0$$

Let $f_n(x) = nx/(1+n^2x^2)$, First, we look at a graph to gain insight. Figure 3.10 shows f_n for $n = 1, 3, 6, 10$. The graphs suggest f_n is bounded by $g(x) = 1/2$. A little elementary calculus verifies this bound.

$$f_n'(x) = \frac{n\left(1 - n^2x^2\right)}{\left(1+n^2x^2\right)^2}$$

So f_n' is zero when $x = 1/n$. Then $f_n(1/n) = 1/2$ shows domination by $g(x) = 1/2$. *Does the bounded convergence theorem also apply?*

By Lebesgue's dominated convergence theorem,

$$\lim_{n\to\infty} \int_0^1 \frac{nx}{1+n^2x^2}\,d\mu = \int_0^1 \left(\lim_{n\to\infty} \frac{nx}{1+n^2x^2}\right)\,d\mu$$
$$= \int_0^1 0\,d\mu = 0$$

Verify that $\lim_{n\to\infty} nx/(1+n^2x^2) = 0$ *for all* $x \in [0,1]$*!* ∎

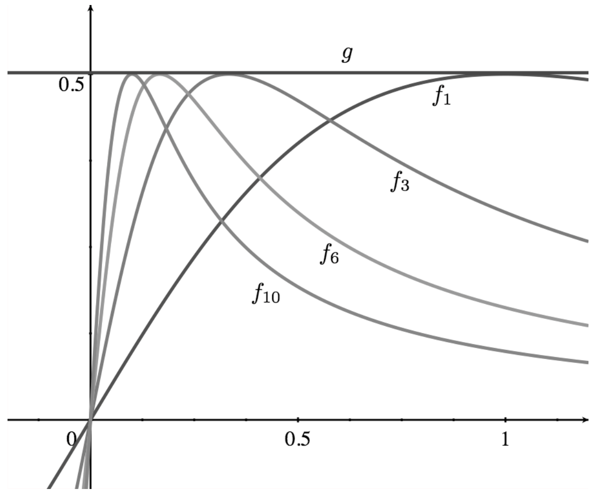


Figure 3.10 Plot of $f_n(x) = nx/(1 + n^2x^2)$ for Several n

Let's also express Lebesgue's dominated convergence theorem in terms of series.

Corollary 3.37 *Let $\{f_n\}$ be a sequence of integrable functions such that $\sum_n f_n$ converges a.e. on $[a, b]$. If there is an integrable function g on $[a, b]$ such that*

$$\left|\sum_{k=1}^{N} f_k(x)\right| \leq g(x) \text{ a.e.}$$

for all N, then series $\sum_n f_n$ is integrable on $[a, b]$, and

$$\int_a^b \left(\sum_{k=1}^{\infty} f_k\right) d\mu = \sum_{k=1}^{\infty} \left(\int_a^b f_k \, d\mu\right)$$

Lebesgue has shown that we can replace uniform convergence with measurability and a dominating function yet still interchange limits and integrals. This very powerful theorem helped to answer many questions in diverse areas of analysis.

We end our study of Lebesgue's theory by stating the last theorem from Lebesgue's text *Leçons sur l'intégration* that uses the power of Lebesgue integration and convergence in a result for elementary calculus.

Theorem 3.38 *A monotonic function has a derivative almost everywhere.*

Look to Boas (1981, p. 155) for a proof.

In the next section, we ask how strange are these new concepts of measure and convergence really?

3.4 LITTLEWOOD'S THREE PRINCIPLES

John E. Littlewood (1885–1977) was a well-respected, highly prolific, classical analyst, which means his research was predominantly in real and complex analysis, as opposed to functional analysis; however, he also worked in other areas such as analytic number theory and differential equations. In 1944, Littlewood wrote a very influential textbook *Lectures on the Theory of Functions* (Littlewood, 1944). Rado says in his review of Littlewood's text for the December 1945 issue of the *The Mathematical Gazette*,

> [The] sole aim [of the text] seems to be to make the reader share the author's delight in one of the most beautiful realms of mathematics.

One of the most referenced heuristics comes from Littlewood's text. He proposed three principles as guides for working in real analysis:

1. Every measurable set is nearly a finite union of intervals.
2. Every measurable function is nearly continuous.
3. Every convergent sequence of measurable functions is nearly uniformly convergent.

Royden (1988, p. 71), among many others, quotes Littlewood (1944, p.26) directly

> Most of the results are fairly intuitive applications of these ideas. If one of the principles would be the obvious means to settle the problem if it were 'quite' true, it is natural to ask if the 'nearly' is near enough, and for a problem that is actually solvable it generally is.

Littlewood's principles correspond to important theorems of Lebesgue (1902), Luzin[1] (1912), and Egorov (1911). However, the point of these maxims is not to highlight the theorems but rather to offer an approach to solving problems in real analysis. For example, suppose we are trying to establish a property that would be easy if the functions involved were continuous. Then we should attempt to prove the result using "nearly continuous," or measurable as measurable is "continuous except on a set of size less than ϵ." *By which theorem?* Many times, this "nearly" true is enough to prove the desired result. Littlewood's principles provide a very important approach to real analysis, so useful that, as in Royden (1988, Section 3.6) or Bichteler (1998, Section III.1), the principles have their own section.

Summary

We have looked at how arbitrary sets can be measured, generalizing the length of an interval. We then learned how to build an integral for successively larger classes of functions. The technique of creating a chain from simple functions to nonnegative to general functions is very useful and appears in many places in analysis. We then examined how much convergence properties were enhanced by Lebesgue's integral,

[1] Actually, in 1903, Lebesgue stated the result named for Luzin without proof; Vitali proved it in 1905.

essentially replacing the need for uniform convergence. The last section discussed how usually the properties we need are nearly true, so the theory isn't nearly as difficult as it first seems.

Today, there are two standard approaches to Lebesgue theory. The one we have used to develop measure is Carathéodory's method of 1914 that asks whether a set splits all other sets into pieces whose sizes add properly. For deeper and more general treatments, see, for example, Cohn (1980), Royden (1988), or Rudin (1976), and especially Bressoud's (2008) *A Radical Approach to Lebesgue's Theory of Integration,* which also shows Lebesgue's technique using inner and outer measures. Lebesgue's method is also illustrated in Boas's (1981) *A Primer of Real Functions*. The other commonly used approach first develops the Lebesgue integral and then defines the measure of a set to be the value of the integral of that set's characteristic function. This method is done very well in Chae's (1995) *Lebesgue Integration*. Burk's (2007) *A Garden of Integrals* is an excellent comparison of several different types of integrals.

There are numerous directions for further study. We could define an indefinite Lebesgue integral and investigate its properties. We could study spaces other than the real numbers to develop more general measure theory. We could generalize to infinite dimensional spaces of functions. We could return to Fourier series and convergence. There are exciting topics in every direction we look.

EXERCISES

3.1 Prove De Morgan's laws for sets.

a) $(A \cup B)^c = A^c \cap B^c$

b) $(A \cap B)^c = A^c \cup B^c$

3.2 Let M be a collection of sets and let

$$\mathcal{M} = \bigcap \{\mathcal{B} \text{ is an algebra and } M \subset \mathcal{B}\}$$

Prove: $\mathcal{M}$ is the smallest algebra containing M.

3.3 Show that the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ also contains all the closed sets.

3.4 Prove the *monotonicity* of Lebesgue measure: if $A \subseteq B \in \mathfrak{M}$, then $\mu(A) \leq \mu(B)$.

3.5 If there is a set in $E \in \mathfrak{M}$ with finite measure, then $\mu(\emptyset) = 0$.

3.6 Let $X = [0, 1]$. Set $\mathcal{A}$ to be the collection of subsets $S \subseteq [0, 1]$ where either S or $S^c = [0, 1] - S$ is finite. Show

a) $\mathcal{A}$ is an algebra.

b) $\mathcal{A}$ is not a σ-algebra.

3.7 Show

a) $\mu^*(\emptyset) = 0$

b) $\mu^*(\{x\}) = 0$

3.8 Prove: μ^* is translation invariant. That is, $\mu^*(x + E) = \mu^*(E)$ for every $E \subset \mathbb{R}$ and $x \in \mathbb{R}$.

3.9 Prove: If E_1 and E_2 are measurable, then $E_1 \cap E_2$ is measurable.

3.10 Prove Theorem 3.13.

3.11 Suppose that $E \in \mathfrak{M}$. Show there exists an open set $O \supseteq E$ and a closed set $F \subseteq E$ such that $\mu(E - F) < \epsilon$ for any given $\epsilon > 0$.

3.12 Suppose $A \subset E$ and $\mu(E) = 0$. Then A is measurable and $\mu(A) = 0$.

3.13 Let A and B be two bounded, measurable subsets of $\mathbb{R}$. Prove

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

3.14 Give a class presentation on the *inclusion-exclusion principle*.

3.15 Prove Corollary 3.16, if a function satisfies the measurability conditions, then the set $f^{-1}(a)$ is measurable for every $a \in \mathbb{R}$.

3.16 Use Theorem 3.17 to prove that all polynomials are measurable.

3.17 Prove the statements of Theorem 3.18 for infima.

3.18 If f is measurable, then $|f| = \max\{f, -f\}$ is measurable. Give an example showing that the measurability of $|f|$ does *not* imply that of f.

3.19 Give an example to show that the measurability of the sets $\{f^{-1}(a)\}$ for every a does not imply that f is a measurable function.

3.20 If $a \leq b$, show that

$$\int \chi_{[a,b]}\, d\mu = b - a$$

3.21 If E is a bounded measurable set, then

$$\int_E d\mu = \int \chi_E\, d\mu = \mu(E)$$

3.22 Prove Theorem 3.26, the properties of Lebesgue integrals.

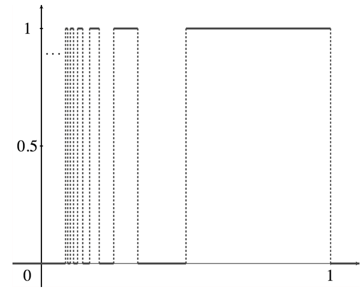
3.23 Suppose f is a bounded measurable function and E is a set of measure zero. Find $\int_E f\, d\mu$.

3.24 Prove Corollary 3.27,

$$\left|\int_E f\, d\mu\right| \leq \int_E |f|\, d\mu$$

3.25 Let

$$f(x) = \begin{cases} 1 & \dfrac{1}{2n} < x \leq \dfrac{1}{2n-1}, n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

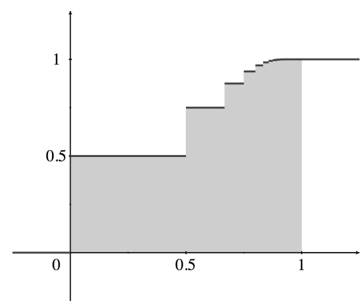


Show that both the Riemann and Lebesgue integrals exist and

$$\int_0^1 f(x)\, dx = \int_0^1 f\, d\mu = \ln(2)$$

3.26 Let $H(x)$ be the *unit step function*, $H(x) = 1$ if $x > 0$ and $H(x) = 0$ otherwise. Define *Zeno's staircase* by

$$Z(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot H\left(x - \frac{k-1}{k}\right)$$



a) Determine if Z is Lebesgue integrable on $[0, 1]$. If so, find the value.

b) Determine if Z is Riemann integrable on $[0, 1]$. If so, find the value.

3.27 Let $S(x) = x^{-1/2}$ for $x \neq 0$ and $S(0) = 0$. Use n-truncations to show that $S \in \mathcal{L}([0, 1])$ to find the value of $\int_0^1 S\, d\mu$ as a limit.

3.28 Prove Theorem 3.30, the algebra of Lebesgue integrals for unbounded functions.

3.29 Suppose f is integrable on $[a, b]$ and $\epsilon > 0$. Prove there is a $\delta > 0$ such that for any subset $E \subset [a, b]$ with $\mu(E) < \delta$ we have

$$\left| \int_E f \, d\mu \right| < \epsilon$$

3.30 Determine whether

$$g_n(x) = \begin{cases} 1 & \dfrac{1}{2n} < x \leq \dfrac{1}{2n-1} \\ 0 & \text{otherwise} \end{cases}$$

converges

a) pointwise on $[0, 1]$.
b) almost everywhere on $[0, 1]$.
c) in mean on $[0, 1]$.

3.31 Discuss the convergence properties of

$$D_n(x) = \begin{cases} n & x \in \mathbb{Q} \\ \dfrac{x^n}{n} & \text{otherwise} \end{cases}$$

for $x \in [0, 1]$.

3.32 Verify the claims in Example 3.11, part 1.

3.33 Verify the claims in Example 3.11, part 2.

3.34 Give an example of a sequence of functions that converges almost uniformly but not pointwise on $\mathbb{R}$.

3.35 Use the characterization of measurable functions as limits almost everywhere of step functions and Egorov's theorem to prove Theorem 3.21, Luzin's theorem.

3.36 Suppose that f is Lebesgue integrable and $|f(x)| < 1$ for all $x \in [-1, 1]$. Show that

$$\int_{[-1,1]} f^n \, d\mu = 0$$

3.37 Determine the value of

$$\int_0^1 \left(\sum_{n=1}^{\infty} \frac{x^n}{n} \right) d\mu$$

3.38 Calculate the value of

$$\int_0^1 \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) d\mu$$

with two different methods. First, using Lebesgue's dominated convergence theorem and, second, by recognizing the summation as a standard function.

3.39 Verify the claims of Example 3.13.

3.40 Define h_n by $h_n(x) = nxe^{-nx^2}$.

a) Compute

$$\int_0^1 \lim_{n \to \infty} h_n \, d\mu$$

Hint: look at graphs of h_n.

b) Compute

$$\lim_{n \to \infty} \int_0^1 h_n \, d\mu$$

c) Explain the results in relation to Lebesgue's dominated convergence theorem.

3.41 Let $f_n(x) = (n+1)x^n$ on $[0, 1]$.

a) Compute

$$\int_{[0,1]} \liminf_{n \to \infty} f_n \, d\mu$$

b) Compute

$$\liminf_{n \to \infty} \int_{[0,1]} f_n \, d\mu$$

c) Explain the results in relation to Fatou's lemma.

3.42 Give a class presentation on Lebesgue's life and mathematics.

3.43 State the three theorems corresponding to Littlewood's three principles.

3.44 Give a class presentation on Littlewood's life and mathematics.

3.45 Who said,

In my opinion, a mathematician, in so far as he is a mathematician, need not preoccupy himself with philosophy—an opinion, moreover, which has been expressed by many philosophers.

APPENDIX A

DEFINITIONS AND THEOREMS OF ELEMENTARY REAL ANALYSIS

The main definitions and results of elementary real analysis are collected here for reference. The theorems that underlie first-year calculus appear along with a number of interesting propositions that provide more depth and insight.

A.1 LIMITS

Definition A.1 (Accumulation Point) *Let $D \subseteq \mathbb{R}$. A point $a \in \mathbb{R}$ is an* accumulation point of D *iff every open interval containing a also contains a point $x \in D$ with $x \neq a$.*

Definition A.2 *Let $f : D \to \mathbb{R}$ and a be an accumulation point of D. Then*

$$\lim_{x \to a} f(x) = L$$

iff for every $\epsilon > 0$ there is a $\delta > 0$ such that whenever $x \in D$ and $0 < |x - a| < \delta$ then $|f(x) - L| < \epsilon$.

Copyright © 2009 John Wiley & Sons, Inc.

Theorem A.1 (Algebra of Limits) *Suppose that $f, g : D \to \mathbb{R}$ both have finite limits at $x = a$ an accumulation point of D and let $c \in \mathbb{R}$. Then*

- $\lim\limits_{x \to a} c\, f(x) = c \lim\limits_{x \to a} f(x)$
- $\lim\limits_{x \to a} f(x) \pm g(x) = \lim\limits_{x \to a} f(x) \pm \lim\limits_{x \to a} g(x)$
- $\lim\limits_{x \to a} f(x) \cdot g(x) = \lim\limits_{x \to a} f(x) \cdot \lim\limits_{x \to a} g(x)$
- *if* $\lim\limits_{x \to a} g(x) \neq 0$, *then* $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim\limits_{x \to a} f(x)}{\lim\limits_{x \to a} g(x)}$

Theorem A.2 ("Sandwich Theorem") *Suppose that $g(x) \leq f(x) \leq h(x)$ for all $x \in (a - h, a + h)$ for some $h > 0$. If $\lim_{x \to a} g(x) = L = \lim_{x \to a} h(x)$, then $\lim_{x \to a} f(x) = L$.*

A.2 CONTINUITY

Definition A.3 *Let $f : D \to \mathbb{R}$ and $a \in D$. Then f* is continuous at *$x = a$ iff a is an isolated point of D or for every $\epsilon > 0$ there is a $\delta > 0$ such that whenever $x \in D$ and $|x - a| < \delta$ then $|f(x) - f(a)| < \epsilon$.*

Theorem A.3 (Algebra of Continuity) *Suppose that $f, g : D \to \mathbb{R}$ both are continuous at $x = a \in D$ and that $c \in \mathbb{R}$. Then*

- *cf is continuous at a*
- *$f \pm g$ is continuous at a*
- *$f \cdot g$ is continuous at a*
- *if $g(a) \neq 0$, then f/g is continuous at a*

Corollary A.4 *Every real polynomial is continuous at every $x \in \mathbb{R}$.*

Theorem A.5 *If a function f is continuous at a and ϕ is a function such that $\lim_{t \to t_0} \phi(t) = a$, then*

$$\lim_{t \to t_0} f\left(\phi(t)\right) = f\left(\lim_{t \to t_0} \phi(t)\right)$$

Theorem A.6 (Continuity of Composition) *Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ where $f(A) \subseteq B$. Suppose that f is continuous at $x = a \in A$, that g is continuous at $x = f(a) \in B$. Then $g \circ f$ is continuous at $x = a$.*

Theorem A.7 *If a function f is continuous on a closed, bounded interval $[a, b]$, then f is bounded on $[a, b]$.*

Theorem A.8 (Intermediate Value Theorem) *If $f : [a, b] \to \mathbb{R}$ is continuous and if k is between $f(a)$ and $f(b)$, then there exists $c \in (a, b)$ such that $f(c) = k$.*

Corollary A.9 *Every odd-degree real polynomial has a real root.*

Corollary A.10 *Every real polynomial is a product of linear factors and irreducible (over $\mathbb{R}$) quadratic factors.*

Corollary A.11 (Fundamental Theorem of Algebra) *Every nth-degree real polynomial has n complex roots counting multiplicity.*

Theorem A.12 (Extreme Value Theorem) *If $f : [a, b] \to \mathbb{R}$ is continuous, then*

1. *there exists $x_m \in [a, b]$ such that $f(x_m) = \min_{x \in [a,b]} f(x)$*
2. *there exists $x_M \in [a, b]$ such that $f(x_M) = \max_{x \in [a,b]} f(x)$*

Definition A.4 (Uniform Continuity) *A function $f : D \to \mathbb{R}$ is* uniformly continuous *on $E \subseteq D$ iff for every $\epsilon > 0$ there is a $\delta > 0$ such that whenever $x_1, x_2 \in E$ and $|x_1 - x_2| < \delta$, then $|f(x_1) - f(x_2)| < \epsilon$.*

Theorem A.13 *If f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.*

Definition A.5 (Discontinuities) *If a function $f : D \to \mathbb{R}$ is not continuous at a point $a \in D$ and both $f(a+) = \lim_{x \to a+} f(x)$ and $f(a-) = \lim_{x \to a-} f(x)$ exist, then $x = a$ is a* simple discontinuity *of f. Further, if $f(a+) = f(a-) \neq f(a)$, then $x = a$ is a* removable discontinuity*; if $f(a+) \neq f(a-)$, then $x = a$ is a* jump discontinuity.

A point of discontinuity that is not simple [either $f(a+)$ or $f(a-)$ fails to exist] is an essential discontinuity.

Theorem A.14 *If $f : [a, b] \to \mathbb{R}$ is monotonic, then the set of points of discontinuity is countable.*

Theorem A.15 *The set of discontinuities of any function must be a countable union of closed sets or an F_σ set.*

Corollary A.16 *A function cannot be discontinuous only on the irrational reals.*

A.3 THE DERIVATIVE

Definition A.6 *Let $f : D \to \mathbb{R}$ and $a \in D$ be an accumulation point. Then*

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

Theorem A.17 *If f is differentiable at $x = a$, then f is continuous at $x = a$.*

Theorem A.18 (Algebra of Derivatives) *If $f, g : D \to \mathbb{R}$ are differentiable at $x = a$ and $c \in \mathbb{R}$, then at $x = a$,*

- $(cf)' = c(f')$
- $(f \pm g)' = f' \pm g'$
- $(f \cdot g)' = f' \cdot g + f \cdot g'$
- *if* $g(a) \neq 0$, *then* $\left(\dfrac{f}{g}\right)' = \dfrac{f' \cdot g + f \cdot g'}{g^2}$

Theorem A.19 (The Chain Rule) *Let $f : A \to B$ and $g : B \to \mathbb{R}$. Suppose that f is differentiable at $x = a \in A$ and that g is differentiable at $x = b = f(a) \in B$. Then $g \circ f$ is differentiable at $x = a$ and*

$$(g \circ f)'(a) = g'\left(f(a)\right) \cdot f'(a)$$

Corollary A.20 *Let u be a differentiable function of x and $r \in \mathbb{R}$. Then, when defined,*

$$(u^r)' = r\, u^{r-1} \cdot u'$$

$$(e^u)' = e^u \cdot u' \qquad \ln(u)' = \frac{1}{u} \cdot u'$$

$$\sin(u)' = \cos(u) \cdot u' \qquad \cos(u)' = -\sin(u) \cdot u'$$

$$\tan(u)' = \sec^2(u) \cdot u' \qquad \cot(u)' = -\csc^2(u) \cdot u'$$

$$\sec(u)' = \sec(u)\tan(u) \cdot u' \qquad \csc(u)' = -\csc(u)\cot(u) \cdot u'$$

$$\sin^{-1}(u)' = \frac{1}{\sqrt{1-u^2}} \cdot u' \qquad \cos^{-1}(u)' = \frac{-1}{\sqrt{1-u^2}} \cdot u'$$

$$\tan^{-1}(u)' = \frac{1}{1+u^2} \cdot u' \qquad \cot^{-1}(u)' = \frac{-1}{1+u^2} \cdot u'$$

$$\sec^{-1}(u)' = \frac{1}{|u|\sqrt{u^2-1}} \cdot u' \qquad \csc^{-1}(u)' = \frac{-1}{|u|\sqrt{u^2-1}} \cdot u'$$

Theorem A.21 (Inverse Function Theorem) *Let $f : [a, b] \to \mathbb{R}$ be differentiable with $f'(x) \neq 0$ for any $x \in [a, b]$. Then*

- *f is injective (1–1)*
- *f^{-1} is continuous on $f([a, b])$*
- *f^{-1} is differentiable on $f([a, b])$*
- $(f^{-1})'(y) = \dfrac{1}{f'(x)}$ *where $y = f(x)$*

Theorem A.22 *Suppose that $f : [a, b] \to \mathbb{R}$ has an extremum at $c \in (a, b)$. If f is differentiable at $c \in (a, b)$, then $f'(c) = 0$.*

Theorem A.23 (Rolle's Theorem) *If $f : D \to \mathbb{R}$ is continuous on $[a, b] \subseteq D$ and differentiable on (a, b) with $f(a) = f(b)$, then there exists a value $c \in (a, b)$ such that $f'(c) = 0$.*

Theorem A.24 (Mean Value Theorem) *If $f : D \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists a value $c \in (a, b)$ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

Corollary A.25 *If $f : D \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists a value $\theta \in (0, 1)$ such that*

$$f(a + h) = f(a) + h \cdot f'(a + \theta h)$$

Corollary A.26 (Racetrack Principle) *If $f : D \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b), and $m \leq f'(x) \leq M$, then*

$$f(a) + m \cdot (b - a) \leq f(b) \leq f(a) + M \cdot (b - a)$$

Corollary A.27 *If $f : D \to \mathbb{R}$ is continuous on $[a, b]$, differentiable on (a, b), and $f'(x) = 0$ on D, then f is a constant function.*

Corollary A.28 *If $f, g : D \to \mathbb{R}$ are continuous on $[a, b]$, differentiable on (a, b), and $f'(x) = g'(x)$ on D, then $f(x) = g(x) + k$ on D where k is a constant.*

Corollary A.29 *If $f : D \to \mathbb{R}$ is differentiable on $[a, b]$, then f' has the* intermediate value property.

Theorem A.30 (Cauchy's Mean Value Theorem) *If $f, g : D \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b), then there exists a value $c \in (a, b)$ such that*

$$f'(c)\,[g(b) - g(a)] = g'(c)\,[f(b) - f(a)]$$

or, when denominators are nonzero,

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

Theorem A.31 (Darboux Intermediate Value Theorem) *Let $f : D \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b), from the right at a and the left at b. Then f' has the intermediate value property; i.e., for every t between $f'_{+}(a)$ and $f'_{-}(b)$, there is an $\hat{x} \in [a, b]$ with $f(\hat{x}) = t$.*

Definition A.7 (Uniform Differentiability) *Let $f : [a,b] \to \mathbb{R}$. Then f is* uniformly differentiable on $[a,b]$ *iff f is differentiable on $[a,b]$ and, for every $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $x_1, x_2 \in [a,b]$ with $|x_1 - x_2| < \delta$ it must follow that*

$$\left|\frac{f(x_1) - f(x_2)}{x_1 - x_2} - f'(x_1)\right| < \epsilon$$

Corollary A.32 *If $f : D \to \mathbb{R}$ is uniformly differentiable on $[a,b]$, then f' is continuous on $[a,b]$.*

Definition A.8 (Lipschitz Condition) *Let $f : D \to \mathbb{R}$. If there are positive constants M and α such that for any $x_1, x_2 \in D$*

$$|f(x_1) - f(x_2)| \leq M \cdot |x_1 - x_2|^{\alpha}$$

then f is Lipschitz-α *with constant M, written $f \in \mathrm{Lip}_M\, \alpha$.*

Theorem A.33 *If $f \in \mathrm{Lip}_M\, \alpha$ on D, then*

1. *f is continuous,*
2. *if $\alpha > 1$, f is constant.*

Corollary A.34 *If $f : [a,b] \to \mathbb{R}$ is differentiable, then $f \in \mathrm{Lip}_M\, 1$.*

Definition A.9 (Measure Zero) *A set E has* measure zero *if and only if for any $\epsilon > 0$ the set E can be covered by a countable collection of open intervals having total length less than ϵ; i.e., $E \subseteq \bigcup_i (a_i, b_i)$ where $\sum_i (b_i - a_i) < \epsilon$.*

Definition A.10 (Almost Everywhere) *A property P holds* almost everywhere *if the set $\{x : P(x)$ is not true$\}$ has measure zero.*

Theorem A.35 (Rademacher's Theorem) *If $f \in \mathrm{Lip}_M\, 1$, then f is differentiable almost everywhere.*

Theorem A.36 (Lebesgue Differentiation Theorem) *If $f : \mathbb{R} \to \mathbb{R}$ is continuous and nondecreasing, then f is differentiable almost everywhere.*

Definition A.11 (Higher Order Derivatives) *The nth derivative of $f(x)$, if it exists, is given by*

$$f^{(n)}(x) = \frac{d}{dx} f^{(n-1)}(x)$$

for $n > 1$ where $f^{(0)} = f$.

Theorem A.37 *Let $f : D \to \mathbb{R}$ be m times continuously differentiable. Then f has* a root of multiplicity m at $x = r$ *iff $f^{(m)}(r) \neq 0$, but*

$$f(r) = f'(r) = \cdots = f^{(m-1)}(r) = 0$$

Theorem A.38 (First Derivative Test for Extrema) *Let f be continuous on $[a, b]$ and differentiable on (a, b). If $f'(c) = 0$ for $c \in (a, b)$ and f' changes sign from*

- *negative to positive around c, then c is a relative minimum of f;*
- *positive to negative around c, then c is a relative maximum of f.*

If f' does not change sign around c, then c is a stationary "terrace point" of f.

Theorem A.39 (Second Derivative Test for Extrema) *Let f be continuous on $[a, b]$ and twice differentiable on (a, b). If $f'(c) = 0$ for $c \in (a, b)$ and*

- *$f''(c)$ is positive, then c is a relative minimum of f;*
- *$f''(c)$ is negative, then c is a relative maximum of f;*
- *$f''(c) = 0$, then the test fails.*

Theorem A.40 (Taylor's Theorem or Extended Law of the Mean) *Let $n \in \mathbb{N}$ and suppose that f has $n + 1$ derivatives on $(a - h, a + h)$ for some $h > 0$. Then for $x \in (a - h, a + h)$*

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} (x - a)^k + R_n(x)$$

where

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n + 1)!} (x - a)^{n+1}$$

for some c between x and a.

Theorem A.41 (Taylor's Theorem with Lagrange's Form of the Remainder) *Let $n \in \mathbb{N}$ and suppose that f has $n + 1$ continuous derivatives on $(a - h, a + h)$ for some $h > 0$. Then for $x \in (a - h, a + h)$*

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} (x - a)^k + L_n(x)$$

where

$$L_n(x) = \frac{1}{n!} \int_a^x f^{(n+1)}(t) \cdot (x - t)^n \, dt$$

Theorem A.42 (Bernstein) *Let I be an interval and $f : I \to \mathbb{R}$ have derivatives of all orders. If f and all its derivatives are nonnegative, then the Taylor series of f converges on I.*

Theorem A.43 (L'Hôpital's Rule) *Suppose that f and g are differentiable on an open interval I containing a and that*

$$\lim_{x \to a} f(x) = 0 = \lim_{x \to a} g(x)$$

while $g'(x) \neq 0$ on I. Then, if the limit exists,

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = \lim_{x\to a} \frac{f(x)}{g(x)}$$

Corollary A.44 *Let $n \in \mathbb{N}$. Then*

$$\lim_{x\to\infty} \frac{x^n}{e^x} = 0 \quad \textit{and} \quad \lim_{x\to\infty} \frac{\ln(x)}{\sqrt[n]{x}} = 0$$

Corollary A.45 *If f is twice differentiable on an open interval I and $x \in I$, then*

$$f''(x) = \lim_{h\to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

A.4 RIEMANN INTEGRATION

Definition A.12 (Partition) *A* partition *P of a closed interval $[a,b]$ is an ordered set of values $\{x_i \,|\, i = 0, \ldots, n\}$ such that $a = x_0 < x_1 < \cdots < x_n = b$. The* norm *or* mesh *of the partition is*

$$\|P\| = \max\{\Delta x_i \,|\, i = 1 \ldots n\}$$

where $\Delta x_i = x_i - x_{i-1}$.

Definition A.13 (Cauchy Sum) *Let $f : [a,b] \to \mathbb{R}$ be a continuous function and P be a partition of $[a,b]$. Then the Cauchy sum (or "left endpoint sum") of f (w.r.t. P) is*

$$C(P,f) = \sum_{k=1}^{n} f(x_{k-1})\Delta x_k$$

Definition A.14 (Riemann Sum) *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, P be a partition of $[a,b]$, and $c_k \in [x_{k-1}, x_k]$ for each k. Then the Riemann sum of f (w.r.t. P and $\{c_k\}$) is*

$$R(P,f) = \sum_{k=1}^{n} f(c_k)\Delta x_k$$

Definition A.15 (Darboux Sums) *Let $f : [a,b] \to \mathbb{R}$ and P be a partition of $[a,b]$ and set*

$$M_k(f) = \sup_{[x_{k-1},x_k]} f(x) \qquad \textit{and} \qquad m_k(f) = \inf_{[x_{k-1},x_k]} f(x)$$

Then the upper and lower Darboux sums of f (w.r.t. P) are

$$U(P,f) = \sum_{k=1}^{n} M_k \Delta x_k \qquad \textit{and} \qquad L(P,f) = \sum_{k=1}^{n} m_k \Delta x_k$$

respectively.

Lemma A.46 *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, say, below by m and above by M, and let P be a partition of $[a,b]$. Then*

$$m(b-a) \leq L(P,f) \leq R(P,f) \leq U(P,f) \leq M(b-a)$$

for all choices of $\{c_k\}$.

Lemma A.47 *Suppose $f : [a,b] \to \mathbb{R}$ is a bounded function and P and Q are partitions of $[a,b]$. If $P \subseteq Q$ (i.e., Q is a* finer *partition), then*

1. $L(P,f) \leq L(Q,f)$ *and* $U(Q,f) \leq U(P,f)$
2. $L(P,f) \leq U(Q,f)$

Definition A.16 *Set*

$$\overline{\int_a^b} f(x)\,dx = \inf_P U(P,f)$$

and

$$\underline{\int_a^b} f(x)\,dx = \sup_P L(P,f)$$

Lemma A.48 *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and P be a partition of $[a,b]$. Then*

$$L(P,f) \leq \underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx \leq U(P,f)$$

Definition A.17 *A bounded function $f : [a,b] \to \mathbb{R}$ is* Riemann integrable *on $[a,b]$ if and only if*

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx = A \in \mathbb{R}$$

Set $\mathfrak{R}[a,b] = \{f \mid f$ is Riemann integrable on $[a,b]\}$.

Theorem A.49 *The bounded function $f : [a,b] \to \mathbb{R}$ is Riemann integrable on $[a,b]$ if and only if for any $\epsilon > 0$ there is a partition P such that $U(P,f) - L(P,f) < \epsilon$.*

Theorem A.50 *If f is monotone on $[a,b]$, then $f \in \mathfrak{R}[a,b]$.*

Theorem A.51 *If f is continuous on $[a,b]$, then $f \in \mathfrak{R}[a,b]$.*

Theorem A.52 *If $f, g \in \mathfrak{R}[a,b]$ and $c \in \mathbb{R}$, then*

1. $f + g \in \mathfrak{R}[a,b]$ *and* $\displaystyle\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$
2. $cf \in \mathfrak{R}[a,b]$ *and* $\displaystyle\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$

Theorem A.53 *If f, $g \in \mathfrak{R}[a,b]$ and $f(x) \leq g(x)$ on $[a,b]$, then*

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$$

Theorem A.54 *If $f \in \mathfrak{R}[a,b]$, then $|f| \in \mathfrak{R}[a,b]$ and*

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

Theorem A.55 *Let $f : [a,b] \to \mathbb{R}$ is bounded and $c \in (a,b)$. Then $f \in \mathfrak{R}[a,b]$ if and only if $f \in \mathfrak{R}[a,c]$ and $f \in \mathfrak{R}[c,b]$ and further*

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

Theorem A.56 *If $f \in \mathfrak{R}[a,b]$ and g is continuous on $f([a,b])$, then $g \circ f \in \mathfrak{R}[a,b]$.*

Corollary A.57 *Let $f, g \in \mathfrak{R}[a,b]$ and $n \in \mathbb{N}$. Then*

1. $f^n \in \mathfrak{R}[a,b]$
2. $f \cdot g \in \mathfrak{R}[a,b]$

Lemma A.58 *Let $f \in \mathfrak{R}[a,b]$ and $c, d \in [a,b]$. Then*

1. $\displaystyle\int_c^c f(x)\,dx = 0$
2. $\displaystyle\int_c^d f(x)\,dx = -\int_d^c f(x)\,dx$

Theorem A.59 (Bonnet's Theorem or First Mean Value Theorem for Integrals) *Let f be continuous on $[a,b]$ and $0 \leq g \in \mathfrak{R}[a,b]$. Then there exists a value $c \in (a,b)$ such that*

$$\int_a^b f(x)\,g(x)\,dx = f(c)\int_a^b g(x)\,dx$$

Theorem A.60 (Second Mean Value Theorem for Integrals) *If f is a monotone function on $[a,b]$, then there exists a value $c \in (a,b)$ such that*

$$\int_a^b f(x)\,dx = f(a)(c-a) + f(b)(b-c)$$

Theorem A.61 (The Fundamental Theorem of Calculus)

- *Define $F(x) = \displaystyle\int_a^x f(t)\,dt$. Then F is continuous and, if f is continuous at x_0, then $F'(x_0) = f(x_0)$.*
- *If $F' = f$ on $[a,b]$, then $\displaystyle\int_a^b f(x)\,dx = F(b) - F(a)$.*

Theorem A.62 (Midpoint Rule for Numerical Integration) *Suppose* $f : [a, b] \to \mathbb{R}$ *is integrable and that* $|f''(x)| \le M_2$ *on* $[a, b]$. *For* n *given, set* $x_k = a + k(b-a)/n$ *with* $k = 0, \ldots, n$ *and set* $\bar{y}_k = f\left((x_k + x_{k+1})/2\right)$ *for* $k = 0, \ldots, n-1$. *Then*

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n}\left(\bar{y}_0 + \bar{y}_1 + \cdots + \bar{y}_{n-1}\right)$$

with the absolute value of the error bounded by

$$|\mathit{error}| \le \frac{1}{24}\,\frac{(b-a)^3}{n^2} M_2$$

Theorem A.63 (Trapezoid Rule) *Suppose* $f : [a, b] \to \mathbb{R}$ *is integrable and that* $|f''(x)| \le M_2$ *on* $[a, b]$. *For* n *given, set* $x_k = a + k(b-a)/n$ *and set* $y_k = f(x_k)$ *with* $k = 0, \ldots, n$. *Then*

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2n}\left(y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n\right)$$

with the absolute value of the error bounded by

$$|\mathit{error}| \le \frac{1}{12}\,\frac{(b-a)^3}{n^2} M_2$$

Theorem A.64 (Simpson's Rule) *Suppose* $f : [a, b] \to \mathbb{R}$ *is integrable and that* $\left|f^{(4)}(x)\right| \le M_4$ *on* $[a, b]$. *For* n *even, set* $x_k = a + k(b-a)/n$ *and set* $y_k = f(x_k)$ *with* $k = 0, \ldots, n$. *Then*

$$\int_a^b f(x)\,dx \approx \frac{b-a}{3n}\left(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 4y_{n-1} + y_n\right)$$

with the absolute value of the error bounded by

$$|\mathit{error}| \le \frac{1}{180}\,\frac{(b-a)^5}{n^4} M_4$$

Theorem A.65 (Cauchy-Bunyakovsky-Schwarz Inequality) *If* $f, g \in \mathfrak{R}[a, b]$, *then*

$$\left[\int_a^b f(x)\,g(x)\,dx\right]^2 \le \left[\int_a^b f^2(x)\,dx\right] \cdot \left[\int_a^b g^2(x)\,dx\right]$$

A.5 RIEMANN-STIELTJES INTEGRATION

Definition A.18 *Let* α *be a monotonically increasing function on* $[a, b]$. *For any partition* P *define* $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$.

Definition A.19 (Upper and Lower Riemann-Stieltjes Integrals) *Let f be a function that is bounded and α be monotonically increasing on $[a, b]$. For each partition P, define the* upper *and* lower Riemann-Stieltjes sums *by*

$$U(P, f, \alpha) = \sum_{k=1}^{n} M_k \Delta\alpha_k$$
$$L(P, f, \alpha) = \sum_{k=1}^{n} m_k \Delta\alpha_k$$

Now, define the upper *and* lower Riemann-Stieltjes integrals *as*

$$\overline{\int_a^b} f(x)\, d\alpha(x) = \inf_P U(P, f, \alpha)$$
$$\underline{\int_a^b} f(x)\, d\alpha(x) = \inf_P L(P, f, \alpha)$$

Definition A.20 *If $\overline{\int} f d\alpha = \underline{\int} f d\alpha$, then f is* Riemann-Stieltjes integrable *and we write $f \in \mathfrak{R}(\alpha)$.*

Theorem A.66 *A function f is Riemann-Stieltjes integrable on $[a, b]$ if and only if for every $\epsilon > 0$ there is a partition P of $[a, b]$ such that*

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$

Theorem A.67 *If f is continuous on $[a, b]$, then $f \in \mathfrak{R}(\alpha)$ on $[a, b]$.*

Theorem A.68 *If f is monotonic on $[a, b]$ and α is continuous, then $f \in \mathfrak{R}(\alpha)$ on $[a, b]$.*

Theorem A.69 *If f is bounded and has finitely many discontinuities on $[a, b]$ and α is continuous at each discontinuity of f, then $f \in \mathfrak{R}(\alpha)$.*

Theorem A.70 *If f is bounded on $[a, b]$ and $f \in \mathfrak{R}(\alpha)$, then there exists m and $M \in \mathbb{R}$ such that*

$$m(\alpha(b) - \alpha(a)) \leq \int_a^b f(x)\, d\alpha(x) \leq M(\alpha(b) - \alpha(a))$$

Theorem A.71 *Let f and $g \in \mathfrak{R}(\alpha)$ on $[a, b]$ and $c \in \mathbb{R}$. Then*

1. $\int_a^b cf(x)\, d\alpha(x) = c \int_a^b f(x)\, d\alpha(x)$
2. $\int_a^b f(x)\, d\, c\alpha(x) = c \int_a^b f(x)\, d\alpha(x)$
3. $\int_a^b (f + g)(x)\, d\alpha(x) = \int_a^b f(x)\, d\alpha(x) + \int_a^b g(x)\, d\alpha(x)$
4. $\int_a^b f(x)\, d(\alpha_1(x) + \alpha_2(x)) = \int_a^b f(x)\, d\alpha_1(x) + \int_a^b f(x)\, d\alpha_2(x)$
5. $f \cdot g \in \mathfrak{R}(\alpha)$

Theorem A.72 *Let f and $g \in \mathfrak{R}(\alpha)$ on $[a, b]$. If $f(x) \leq g(x)$, then*

$$\int_a^b f(x)\, d\alpha(x) \leq \int_a^b g(x)\, d\alpha(x)$$

Theorem A.73 *If $f \in \mathfrak{R}(\alpha)$, then*

$$\left| \int_a^b f(x)\, d\alpha(x) \right| \leq \int_a^b \left| f(x) \right|\, d\alpha(x)$$

Definition A.21 *Define the* Heaviside function *to be*

$$U(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Theorem A.74 *If f is bounded on $[a, b]$ and continuous at $x_0 \in (a, b)$, then*

$$\int_a^b f(x)\, dU(x - x_0) = f(x_0)$$

Theorem A.75 *Let $c_n \geq 0$ with $\sum_n c_n$ converging and $\{x_n\}$ be a sequence of distinct points in (a, b). Define*

$$\alpha(x) = \sum_{k=1}^{\infty} c_k U(x - x_k)$$

Let f be continuous on $[a, b]$. Then

$$\int_a^b f(x)\, d\alpha(x) = \sum_{k=1}^{\infty} c_k f(x_k)$$

Theorem A.76 *If f is bounded and $\alpha' \in \mathfrak{R}(\alpha)$, then $f \in \mathfrak{R}(\alpha)$ iff $f\alpha' \in \mathfrak{R}$ and*

$$\int_a^b f(x)\, d\alpha(x) = \int_a^b f(x)\alpha'(x)\, dx$$

Theorem A.77 (Hölder's Inequality) *Let f and g be in $\mathfrak{R}(\alpha)$ and let $p, q > 0$ be such that $1/p + 1/q = 1$. Then*

$$\left| \int_a^b f(x)g(x)\, d\alpha(x) \right| \leq \left[\int_a^b |f(x)|^p d\alpha(x) \right]^{1/p} \left[\int_a^b |g(x)|^q d\alpha(x) \right]^{1/q}$$

If $p = 2$, this is called the Cauchy-Bunyakovski-Schwarz inequality.

Theorem A.78 (Minkowski's Inequality) *Let $p > 1$ and let f^p and g^p be in $\mathfrak{R}(\alpha)$. Then*

$$\left[\int_a^b [f(x) + g(x)]^p\, d\alpha(x) \right]^{1/p} \leq \left[\int_a^b |f(x)|^p d\alpha(x) \right]^{1/p} + \left[\int_a^b |g(x)|^p d\alpha(x) \right]^{1/p}$$

A.6 SEQUENCES AND SERIES OF CONSTANTS

Definition A.22 *A* real-valued sequence *is a function* $a : \mathbb{N} \to \mathbb{R}$. *Sequence terms are denoted by* $a(n) = a_n$.

Definition A.23 *Let* $f : \mathbb{N} \to \mathbb{N}$ *be a strictly increasing function. The composition* $a \circ f$ *forms a* subsequence *and is denoted by* $a(f(k)) = a_{n_k}$.

Definition A.24 *A sequence* converges, *written* $\lim_{n\to\infty} a_n = L$, *iff for every* $\epsilon > 0$ *there is an* $N \in \mathbb{N}$ *such that, if* $n > N$, *then* $|a_n - L| < \epsilon$.

Theorem A.79 (Corollary to the Heine-Borel Theorem) *A bounded sequence has a convergent subsequence.*

Definition A.25 (Cauchy Sequence) *A sequence* a_n *is* Cauchy *iff, for every* $\epsilon > 0$, *there is an* $N \in \mathbb{N}$ *such that, if* $n, m > N$, *then* $|a_n - a_m| < \epsilon$.

Theorem A.80 *A sequence is Cauchy iff the sequence converges.*

Definition A.26 *A sequence* $\{a_n\}$ *is*

- monotonically increasing *iff* $a_n \leq a_{n+1}$ *for all* n,
- monotonically decreasing *iff* $a_n \geq a_{n+1}$ *for all* n.

Theorem A.81 *If* $\{a_n\}$ *is monotonic, then* $\{a_n\}$ *converges iff it is bounded.*

Definition A.27 *Let* $\{a_n\}$ *be a sequence. Define*

- $\limsup\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty}\left(\sup\limits_{k\geq n} a_k\right)$ *which is also equal to* $\inf\limits_{n\geq 0}\left(\sup\limits_{k\geq n} a_k\right)$
- $\liminf\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty}\left(\inf\limits_{k\geq n} a_k\right)$ *which is also equal to* $\sup\limits_{n\geq 0}\left(\inf\limits_{k\geq n} a_k\right)$

Theorem A.82 *Let* $\{a_n\}$ *be a real-valued sequence. Then* $\lim_{n\to\infty} a_n = a$ *iff* $\liminf_{n\to\infty} a_n = a = \limsup_{n\to\infty} a_n$.

Definition A.28 *Let* $\{a_n\}$ *be a sequence. The associated* series *is* $s_n = \sum_{k=1}^{n} a_n$. *The terms* s_n *are called the* partial sums *of the series. The series converges to* s, *written as* $s = \sum_{k=1}^{\infty} a_n$, *iff the sequence of partial sums* $\{s_n\}$ *converges to* s.

Theorem A.83 *If* $s_n = \sum_{k=1}^{\infty} a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$.

Definition A.29 (Cauchy Series) *A series* $s_n = \sum_{k=1}^{\infty} a_n$ *is called* Cauchy *iff for every* $\epsilon > 0$ *there is an* $N \in \mathbb{N}$ *such that, if* $n \geq m > N$, *then*

$$\left|\sum_{k=m+1}^{n} a_k\right| < \epsilon$$

Theorem A.84 *Every Cauchy series in $\mathbb{R}$ converges.*

Theorem A.85 (The Comparison Test) *Let $\sum_n a_n$ be a series.*

- *If $|a_n| \leq c_n$ for all $n \geq N_0$ and $\sum c_n$ converges, then $\sum a_n$ converges.*
- *If $a_n \geq d_n > 0$ for all $n \geq N_0$ and $\sum d_n$ diverges, then $\sum a_n$ diverges.*

Theorem A.86 (The Limit Comparison Test) *Let $\sum_n a_n$ and $\sum_n b_n$ be positive series. If $0 < \lim_{n\to\infty} a_n/b_n < \infty$, then the series either both converge or both diverge.*

Theorem A.87 (Cauchy Condensation Test) *Let $\{a_n\}$ be a nonnegative decreasing sequence. Then the series $\sum_n a_n$ converges if and only if $\sum_n (2^n a_{2^n})$ converges.*

Theorem A.88 (The Ratio Test, I) *The series $\sum_n a_n$ converges if*

$$\limsup_n \left| \frac{a_{n+1}}{a_n} \right| < 1$$

and diverges if $|a_{n+1}/a_n| > 1$ for $n \geq N_0$.

Corollary A.89 (The Ratio Test, II) *For the series $\sum_n a_n$, define*

$$\rho = \lim_n \left| \frac{a_{n+1}}{a_n} \right|$$

Then,

- *If $\rho < 1$, the series converges.*
- *If $\rho > 1$, the series diverges.*
- *If $\rho = 1$, the test fails.*

Theorem A.90 (The Root Test) *For the series $\sum_n a_n$, define*

$$\rho = \limsup_n \sqrt[n]{|a_n|}$$

Then,

- *If $\rho < 1$, the series converges.*
- *If $\rho > 1$, the series diverges.*
- *If $\rho = 1$, the test fails.*

Theorem A.91 (The Integral Test) *Let $f : [1, \infty) \to \mathbb{R}$ be a continuous, positive, decreasing function such that $f(n) = a_n$. Then the improper integral $\int_1^\infty f(x)\,dx$ converges if and only if the series $\sum_n a_n$ converges.*

Theorem A.92 (Alternating Series Test) *If $a_n \geq a_{n+1} > 0$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} a_n = 0$, then $\sum_n (-1)^n a_n$ converges.*

A.7 SEQUENCES AND SERIES OF FUNCTIONS

Definition A.30 (Pointwise Convergence) *A sequence of functions* $f_n : E \to \mathbb{R}$ converges pointwise *on* E *iff for each* $x \in E$ *the sequence* $\{f_n(x)\}$ *converges.*

Definition A.31 (Convergence in Mean) *A sequence of integrable real functions* $f_n : [a,b] \to \mathbb{R}$ converges in mean *to* f *iff*

$$\lim_{n\to\infty} \left[\int_a^b [f_n(x) - f(x)]^2 dx \right]^{1/2} = 0$$

Definition A.32 (Uniform Convergence) *A sequence of functions* $f_n : E \to \mathbb{R}$ converges uniformly *to* f *on* E *iff for every* $\epsilon > 0$ *there is an* $N \in \mathbb{N}$ *such that whenever* $n > N$ *then* $|f_n(x) - f(x)| < \epsilon$ *for all* $x \in E$.

Definition A.33 (Cauchy Criterion for Uniform Convergence) *A sequence of functions* $f_n : E \to \mathbb{R}$ *converges uniformly on* E *iff for every* $\epsilon > 0$ *there is an* $N \in \mathbb{N}$ *such that whenever* $n, m > N$ *then* $|f_n(x) - f_m(x)| < \epsilon$ *for all* $x \in E$.

Theorem A.93 *Let* $f_n : E \to \mathbb{R}$ *be a sequence of functions converging uniformly to* f *on* E. *If each* f_n *is continuous on* E, *then* f *is continuous on* E.

Theorem A.94 *Let* $f_n : E \to \mathbb{R}$ *be a sequence of functions converging uniformly to* f *on* E. *Then*

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

for $x, a \in E$.

Theorem A.95 *Let* $f_n : [a,b] \to \mathbb{R}$ *be a sequence of integrable functions converging uniformly on* $[a,b]$. *Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b \lim_{n\to\infty} f_n(x)\,dx$$

Theorem A.96 (Dini's Theorem) *Let* $f_n : [a,b] \to \mathbb{R}$ *be a monotonic sequence of functions converging pointwise to* f *on* $[a,b]$ *where* $-\infty < a < b < \infty$. *Then* f_n *converges uniformly to* f *on* $[a,b]$.

Theorem A.97 *Let* $f_n : [a,b] \to \mathbb{R}$ *be a sequence of continuously differentiable functions converging pointwise to* f *on* $[a,b]$. *If* f_n' *converges uniformly on* $[a,b]$, *then* f_n *converges uniformly to* f *and* f_n' *converges uniformly to the continuous function* f' *on* $[a,b]$.

Theorem A.98 (The Weierstrass M-test) *Let* $\sum_n f_n(x)$ *be a series of functions all defined on* $D \subseteq \mathbb{R}$. *If there is a convergent series of constants* $\sum_n M_n$ *such that for each* n *we have* $|f_n(x)| \le M_n$ *for all* $x \in D$, *then* $\sum_n f_n(x)$ *converges both uniformly and absolutely on* D.

Theorem A.99 (Abel's Uniform Convergence Test) *Let* $f_n : D \to \mathbb{R}$ *be a bounded, monotonically decreasing sequence of functions and* $\sum_n a_n$ *be a convergent series of constants. Then* $\sum_n a_n f_n(x)$ *converges uniformly on* D.

Theorem C.12 (Bounded Monotone Sequence Theorem) *Every bounded monotone sequence of real numbers has a limit.*

Proof: Assume $X = \{x_n\}$ is a bounded, increasing sequence. If X is finite, then the result is easy. Suppose X is infinite. Apply the Bolzano-Weierstrass theorem to find an accumulation point. Show this point is the limit of the sequence. ∎

Complete the circle by now proving

Theorem C.13 (Completeness of the Real Numbers) $\mathbb{R}$ *is complete.*

Proof: Let A be a bounded, nonempty set of real numbers. Let b be any upper bound of A. Choose any point $x_1 \in A$. If x_1 is the supremum, we're done. If not, set $x_2 = (x_1 + b)/2$. If x_2 is the supremum, we're done. Continue this process to create a bounded, infinite sequence. Show the bounded monotone sequence theorem applies. ∎

We have proved: *Completeness of* $\mathbb{R} \implies$ *Heine-Borel Theorem* $\implies$ *Bolzano-Weierstrass Theorem* $\implies$ *Bounded Monotone Sequence Theorem* $\implies$ *Completeness of* $\mathbb{R}$. Since going around the circle of implications shows any result yields any other, the implications are all really 'if and only if.' What does this imply about the four results?

Finish a presentation with an example showing that $\mathbb{Q}$ is *not* complete.

C.7 VITALI'S NONMEASURABLE SET

Giuseppe Vitali was the first to construct a set that was non-Lebesgue measurable. His construction is based on countable additivity, a critical property of measures. Watch for the axiom of choice to appear!

Let $X = [0, 1)$ and define the operation $\oplus : X \to X$ by addition modulo 1. Then $x \oplus y = x + y - \lfloor x + y \rfloor$. (Note: X is then equivalent to the unit circle via $t \mapsto e^{it}$.) Define addition of a set with a number by $A \oplus x = \{a \oplus x \mid a \in A\}$.

Now define the relation $\sim$ on X by

$$x \sim y \text{ if and only if there is a rational } r \text{ such that } |x - y| = r$$

1. Show that for each rational $r \in \mathbb{Q} \cap \mathbb{X}$ we have $r \sim 0$, and so all rationals are equivalent under $\sim$.
2. Prove that $\sim$ is an equivalence relation on $X = [0, 1)$. Let $[x]$ be *the equivalence class of* x, that is, $[x] = \{y \in X \mid y \sim x\}$.
3. Find $[0]$.

Consider $x_1 = \pi/10$ and $x_2 = \pi/30$. Since $x_1 - x_2 = \pi/15 \notin \mathbb{Q}$, then $[x_1] \neq [x_2]$. Now take $x_3 = (\pi + 5)/10$. Since $x_1 - x_3 = 1/2 \in \mathbb{Q}$, we have $x_1 \sim x_3$, so $[x_1] = [x_3]$.

Since $\sim$ is an equivalence relation, it partitions X. Choose one representative h from each equivalence class in the partition of X. (*Axiom of Choice*!) Gather the representative h to form the set H. Consider the collection of these sets $\mathcal{H} = \{H \oplus r\}$ where r ranges over the rationals in X.

4. Determine whether H is countable or uncountable.

5. Verify that $\mathcal{H}$ is a pairwise-disjoint family; i.e., $(H \oplus r_1) \cap (H \oplus r_2) = \emptyset$ for $r_1 \neq r_2$.

6. Prove that
$$X = \bigcup_{r \in \mathbb{Q} \cap X} (H + r)$$

Since Lebesgue measure is translation invariant, $\mu(H \oplus r) = \mu(H)$ for all $r \in \mathbb{Q} \cap X$. Assume that H is Lebesgue measurable, and $\mu(H) = \lambda$. Then, since $\mathcal{H}$ is a countable family of disjoint sets,

$$1 = \mu(X) = \mu\left(\bigcup_{r \in \mathbb{Q} \cap X} (H + r)\right) = \sum_{r \in \mathbb{Q} \cap X} \lambda$$

We have a contradiction: If $\lambda = 0$, then $1 = 0$. Otherwise, if $\lambda > 0$, then $1 = \infty$. Thus H cannot be Lebesgue measurable.

This construction is due to Vitali (1905). *Even though Vitali's paper is written in Italian, enough of the detail is "in mathematics" that we can understand the paper without being able to read Italian. Try it!*

C.8 SOURCES FOR REAL ANALYSIS PROJECTS

A collection of sources for projects for students of real analysis follows.

- Brabenec's *Resources for the Study of Real Analysis* (2006)
- Snow and Weller's *Exploratory Examples for Real Analysis* (2007)
- Kosmala's *A Friendly Introduction to Analysis* (2004): each chapter ends with a collection of student projects.
- Shakarchi and Lang's *Problems and Solutions for Undergraduate Analysis* (1997)
- Aliprantis and Burkinshaw's *Problems in Real Analysis: A Workbook with Solutions* (1999) (advanced real analysis)

C.9 SOURCES FOR PROJECTS FOR CALCULUS STUDENTS

There are many books and websites containing projects and laboratory exercises for calculus classes. A small selection follows.

- Bauldry and Fiedler's *Calculus Projects with Maple* (1996)
- Crannell, LaRose, and Ratliff's *Writing Projects for Mathematics Courses: Crushed Clowns, Cars & Coffee to Go* (2004)
- Gaughan et al's *Student Research Projects in Calculus* (1991)
- Packel and Wagon's *Animating Calculus: Mathematica Notebooks for the Laboratory* (1997)
- Solow and Fink's *Learning by Discovery: A Lab Manual for Calculus* (1993)
- Stroyan's *Projects for Calculus: The Language of Change* (1998)
- Wood's *Calculus Mysteries and Thrillers* (1999)

Many calculus texts offer projects following the chapters.

- *Calculus: Single and Multivariable*, by Hughes-Hallett et al. (2009)
- "Calculus: Modeling and Application," online text by Moore & Smith (2004)
- *Calculus from Graphical, Numerical, and Symbolic Points of View*, by Ostebee & Zorn (2002)
- *Calculus*, by Stewart (2009)

The last word:

Ancora imparo. —*Michelangelo*