Chapter 3

Integration

To remedy deficiencies of Riemann integration that were discussed in Section 1B, in the last chapter we developed measure theory as an extension of the notion of the length of an interval. Having proved the fundamental results about measures, we are now ready to use measures to develop integration with respect to a measure. As we will see, this new method of integration fixes many of the problems with Riemann integration.

Statue in Milan of Italian mathematician Maria Gaetana Agnesi, who in 1748 published one of the first calculus textbooks. A translation of her book into English was published in 1801. In this chapter, we will develop a method of integration more powerful than methods contemplated by the pioneers of calculus.

3A Integration with Respect to a Measure

We will first define the integral of a nonnegative function with respect to a measure. Then by writing a real-valued function as the difference of two nonnegative functions, we will define the integral of a real-valued function with respect to a measure.

The notation introduced below will be useful in our development of integration.

3.1 Definition ***order relation on functions***

Suppose X is a set and $f, g\colon X \to [-\infty, \infty]$ are functions. The notation $g \leq f$ means $g(x) \leq f(x)$ for all $x \in X$.

Integration of Nonnegative Functions

The symbol d *in the expression* $\int f\, d\mu$ *has no meaning, serving only to separate* f *from* μ. *Because the* d *in* $\int f\, d\mu$ *does not represent another object, some mathematicians prefer typesetting an upright* d *in this situation, producing* $\int f\, \mathrm{d}\mu$. *However, the upright* d *looks jarring to some readers who are accustomed to italicized symbols. This book takes the compromise position of using slanted* d *instead of math-mode italicized* d *in integrals.*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. We will denote the integral of an $\mathcal{S}$-measurable function f with respect to μ by $\int f\, d\mu$. Our basic motivation for the definition of this integral is that we want $\int \chi_A\, d\mu$ to equal $\mu(A)$ for all $A \in \mathcal{S}$.

The operation of integration should also be linear. Thus we want

$$\int \Big(\sum_{k=1}^{m} a_k \chi_{A_k}\Big)\, d\mu$$

to equal the sum $\sum_{k=1}^{m} a_k \mu(A_k)$ for all $A_1, \dots, A_m \in S$ and $a_1, \dots, a_m > 0$ (we restrict to positive values of $a_1, \dots, a_m$ at this stage to avoid expressions of the form $\infty - \infty$ and $0 \cdot \infty$).

Suppose $f\colon X \to [0, \infty]$ is an $\mathcal{S}$-measurable function. As a final piece of motivation for the definition of $\int f\, d\mu$, note that f can be approximated from below by functions of the form discussed in the paragraph above (see 2.81 and its proof). Thus we are led to the following definition.

3.2 Definition ***integral of a nonnegative function***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to [0, \infty]$ is a measurable function. The *integral* of f with respect to μ, denoted $\int f\, d\mu$, is defined by

$$\int f\, d\mu = \sup\Big\{ \sum_{k=1}^{m} a_k \mu(A_k) : \sum_{k=1}^{m} a_k \chi_{A_k} \leq f, \text{ where } m \in \mathbf{Z}^+, \\ a_1, \dots, a_m > 0, \text{ and } A_1, \dots, A_m \in \mathcal{S} \Big\}.$$

3.3 Example *integral of 0 is 0*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Let 0 denote the constant function from X to $[0, \infty]$ whose value is 0 at each element of X. Then

$$\int 0 \, d\mu = 0$$

because if $\sum_{k=1}^{m} a_k \chi_{A_k} \leq 0$ with $a_1, \ldots, a_m > 0$, then each $A_k = \emptyset$ and hence $\sum_{k=1}^{m} a_k \mu(A_k) = 0$.

3.4 Example *integral of characteristic function of rational numbers is 0*

If λ is Lebesgue measure on $\mathbf{R}$, then $\int \chi_{\mathbf{Q}} \, d\lambda = 0$ because if $\sum_{k=1}^{m} a_k \chi_{A_k} \leq \chi_{\mathbf{Q}}$ with each $a_k > 0$, then each A_k is a countable set and hence $\lambda(A_k) = 0$ for each k.

Because the integral (with respect to Lebesgue measure) of the characteristic function of $\mathbf{Q}$ is defined and equals 0, even at this early stage in our development of the integral we have fixed one of the deficiencies of Riemann integration.

The definition of $\int f \, d\mu$ given in 3.2 should remind you of the definition of the lower Riemann integral (see 1.7). Indeed, suppose μ is Lebesgue measure on the Borel subsets of an interval $[a, b]$ and $f \colon [a, b] \to \mathbf{R}$ is a bounded function. Then the right side of the equation in 3.2 would be exactly the lower Riemann integral $L(f, [a, b])$ if we required the sets A_k to be intervals (rather than allowing them to be Borel measurable sets).

The lower Riemann integral is not additive, even for bounded nonnegative measurable functions. For example, if $f = \chi_{\mathbf{Q}}$ and $g = \chi_{\mathbf{R} \setminus \mathbf{Q}}$, then

$$L(f, [0,1]) = 0 \quad \text{and} \quad L(g, [0,1]) = 0 \quad \text{but} \quad L(f + g, [0,1]) = 1.$$

In contrast, if λ is Lebesgue measure on the Borel subsets of $[0, 1]$, then

$$\int f \, d\lambda = 0 \quad \text{and} \quad \int g \, d\lambda = 1 \quad \text{and} \quad \int (f + g) \, d\lambda = 1.$$

More generally, we will prove that $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$ for every measure μ and for all nonnegative measurable functions f and g (see 3.21). Thus integration with respect to a measure, even though it is defined similarly to the lower Riemann integral (with the big exception of allowing measurable sets instead of just intervals), has considerably nicer properties than the lower Riemann integral.

3.5 Example *integration with respect to counting measure is summation*

Suppose μ is counting measure on $\mathbf{Z}^+$ and $b_1, b_2, \ldots$ is a sequence of nonnegative numbers. Think of b as the function from $\mathbf{Z}^+$ to $[0, \infty)$ defined by $b(n) = b_n$. Then

$$\int b \, d\mu = \sum_{n=1}^{\infty} b_n,$$

as you should verify.

The representation of a simple function h in the form $\sum_{k=1}^{m} a_k \chi_{A_k}$ is not unique. Requiring the numbers $a_1, \dots, a_m$ to be nonzero and distinct and the sets $A_1, \dots, A_m$ to be nonempty and disjoint does produce a standard representation [take $A_k = h^{-1}(\{a_k\})$, where $a_1, \dots, a_m$ are the distinct nonzero values of h]. The following lemma shows that in the definition of the integral of a nonnegative function, we could have required that the simple function have the standard representation. Doing so would simplify some proofs but complicate other proofs.

3.6 *Useful lemma*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. Suppose $a_1, \dots, a_m, b_1, \dots, b_n \in \mathbf{R}$, and $A_1, \dots, A_m, B_1, \dots, B_n \in \mathcal{S}$ are such that $\sum_{k=1}^{m} a_k \chi_{A_k} = \sum_{j=1}^{n} b_j \chi_{B_j}$. Then

$$\sum_{k=1}^{m} a_k \mu(A_k) = \sum_{j=1}^{n} b_j \mu(B_j).$$

Proof Suppose A_1 and A_2 are not disjoint. Then we can write

3.7 $$a_1 \chi_{A_1} + a_2 \chi_{A_2} = a_1 \chi_{A_1 \setminus A_2} + a_2 \chi_{A_2 \setminus A_1} + (a_1 + a_2) \chi_{A_1 \cap A_2},$$

where the three sets appearing on the right side of the equation above are disjoint.

Now $A_1 = (A_1 \setminus A_2) \cup (A_1 \cap A_2)$ and $A_2 = (A_2 \setminus A_1) \cup (A_1 \cap A_2)$; each of these unions is a disjoint union. Thus $\mu(A_1) = \mu(A_1 \setminus A_2) + \mu(A_1 \cap A_2)$ and $\mu(A_2) = \mu(A_2 \setminus A_1) + \mu(A_1 \cap A_2)$. Hence

$$a_1 \mu(A_1) + a_2 \mu(A_2) = a_1 \mu(A_1 \setminus A_2) + a_2 \mu(A_2 \setminus A_1) + (a_1 + a_2) \mu(A_1 \cap A_2).$$

The equation above, in conjunction with 3.7, shows that if we replace A_1, A_2 by the disjoint sets $A_1 \setminus A_2, A_2 \setminus A_1, A_1 \cap A_2$ and make the appropriate adjustments to the coefficients $a_1, \dots, a_m$, then the value of the sum $\sum_{k=1}^{m} a_k \mu(A_k)$ is unchanged (although m has increased by 1).

Repeating this process with all pairs of subsets among $A_1, \dots, A_m$ that are not disjoint after each step, in a finite number of steps we can convert the initial list $A_1, \dots, A_m$ into a disjoint list of subsets without changing the value of $\sum_{k=1}^{m} a_k \mu(A_k)$.

The next step is to make the numbers $a_1, \dots, a_m$ distinct. This is done by replacing the sets corresponding to each a_k by the union of those sets, and using finite additivity of the measure μ to show that the value of the sum $\sum_{k=1}^{m} a_k \mu(A_k)$ does not change.

Finally, drop any terms for which $a_k = 0$ or $A_k = \varnothing$, getting the standard representation for a simple function. We have now shown that the original value of $\sum_{k=1}^{m} a_k \mu(A_k)$ is equal to the value if we use the standard representation of the simple function $\sum_{k=1}^{m} a_k \chi_{A_k}$. The same procedure can be used with the representation $\sum_{j=1}^{n} b_j \chi_{B_j}$ to show that $\sum_{j=1}^{n} b_j \mu(\chi_{B_j})$ equals what we would get with the standard representation. Hence $\sum_{k=1}^{m} a_k \mu(A_k) = \sum_{j=1}^{n} b_j \mu(B_j)$. ■

Now we can show that the integral of a nonnegative simple function is what we expect it to be.

3.8 *Integral of a linear combination of characteristic functions*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $B_1, \dots, B_n \in \mathcal{S}$, and $b_1, \dots, b_n > 0$. Then

$$\int \Big(\sum_{j=1}^{n} b_j \chi_{B_j} \Big) \, d\mu = \sum_{j=1}^{n} b_j \mu(B_j).$$

Proof Because $\sum_{j=1}^{n} b_j \chi_{B_j} \leq \sum_{j=1}^{n} b_j \chi_{B_j}$, the definition of $\int \Big(\sum_{j=1}^{n} b_j \chi_{B_j} \Big) \, d\mu$ implies that

$$\int \Big(\sum_{j=1}^{n} b_j \chi_{B_j} \Big) \, d\mu \geq \sum_{j=1}^{n} b_j \mu(B_j).$$

To prove an inequality in the other direction, suppose $a_1, \dots, a_m > 0$ and $A_1, \dots, A_m \in \mathcal{S}$ are such that

3.9
$$\sum_{k=1}^{m} a_k \chi_{A_k} \leq \sum_{j=1}^{n} b_j \chi_{B_j}.$$

Then

$$\sum_{j=1}^{n} b_j \chi_{B_j} - \sum_{k=1}^{m} a_k \chi_{A_k} = \sum_{i=1}^{p} c_i \chi_{C_i}$$

for some $c_1, \dots, c_p > 0$ and some $C_1, \dots, C_p \in \mathcal{S}$ (this follows from writing the left side of the equation above in the standard representation for a simple function; 3.9 implies that each $c_i > 0$ in this standard representation).

Now apply 3.6 to both sides of the equation above, getting

$$\sum_{j=1}^{n} b_j \mu(B_j) - \sum_{k=1}^{m} a_k \mu(A_k) = \sum_{i=1}^{p} c_i \mu(C_i).$$

The right side of the equation above is nonnegative. Hence

$$\sum_{k=1}^{m} a_k \mu(A_k) \leq \sum_{j=1}^{n} b_j \mu(B_j).$$

Because $\int \Big(\sum_{j=1}^{n} b_j \chi_{B_j} \Big) \, d\mu$ is defined to be the supremum of values of the left side of the inequality above, we conclude that $\int \Big(\sum_{j=1}^{n} b_j \chi_{B_j} \Big) \, d\mu \leq \sum_{j=1}^{n} b_j \mu(B_j)$, which completes the proof. ∎

Note that 3.8 implies that if $(X, \mathcal{S}, \mu)$ is a measure space and $f \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, then

$$\int f \, d\mu = \sup \Big\{ \int h \, d\mu : h \text{ is a simple } \mathcal{S}\text{-measurable function and } 0 \leq h \leq f \Big\}.$$

As a special case of 3.8, if $(X, \mathcal{S}, \mu)$ is a measure space and $B \in \mathcal{S}$, then $\int \chi_B \, d\mu = \mu(B)$, as we expect.

The next three easy results give unsurprising properties of integrals.

3.10 *Integration is order preserving*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \colon X \to [0, \infty]$ are $\mathcal{S}$-measurable functions such that $f \leq g$. Then $\int f \, d\mu \leq \int g \, d\mu$.

Proof The supremum defining $\int f \, d\mu$ is taken over a subset of the corresponding set for the supremum defining $\int g \, d\mu$. Thus $\int f \, d\mu \leq \int g \, d\mu$. ■

3.11 *Bounding the integral of a nonnegative function*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \colon X \to [0, \infty)$ is $\mathcal{S}$-measurable. Then

$$\int f \, d\mu \leq \mu(X) \sup_{x \in X} f(x).$$

Proof Let $c = \sup_{x \in X} f(x)$. Because $f \leq c$, we have

$$\int f \, d\mu \leq \int c \, d\mu = c\mu(X),$$

where the inequality above comes from 3.10 and the equality above comes from 3.8 (with $n = 1$ and $B_1 = X$). ■

3.12 *Integration is positively homogeneous*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $c \geq 0$. Then $\int cf \, d\mu = c \int f \, d\mu$.

Proof The supremum defining $\int cf \, d\mu$ is taken over a set consisting of c times the set whose supremum defines $\int f \, d\mu$. Thus $\int cf \, d\mu = c \int f \, d\mu$. ■

Proving the additivity of integration is considerably more complicated than proving homogeneity (3.12). The next two results give special cases of additivity (first requiring both functions to be simple, then requiring one of the functions to be constant), which will be used in the next subsection to prove additivity more generally.

3.13 *Integral of a sum of nonnegative simple functions*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \colon X \to [0, \infty)$ are $\mathcal{S}$-measurable simple functions. Then $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.

Proof The desired result follows immediately from 3.8. ■

3.14 *Integral of a constant plus a function*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to [0, \infty)$ is an $\mathcal{S}$-measurable function. Then

$$\int (c + f)\, d\mu = c\mu(X) + \int f\, d\mu$$

for every $c > 0$.

Proof Let $c > 0$. First suppose

3.15 $$a_1, \dots, a_m > 0 \quad \text{and} \quad A_1, \dots, A_m \in \mathcal{S} \quad \text{and} \quad \sum_{k=1}^{m} a_k \chi_{A_k} \leq f.$$

Then $c\chi_X + \sum_{k=1}^{m} a_k \chi_{A_k} \leq c + f$. Thus the definition of the integral of $c + f$ implies

$$\int (c + f)\, d\mu \geq c\mu(X) + \sum_{k=1}^{m} a_k \mu(A_k).$$

In the inequality above, taking the supremum over all choices satisfying 3.15 shows that

$$\int (c + f)\, d\mu \geq c\mu(X) + \int f\, d\mu.$$

To prove the inequality in the other direction, suppose now that

3.16 $a_1, \dots, a_m > 0$ and $A_1, \dots, A_m \in \mathcal{S}$ are disjoint and $\displaystyle\sum_{k=1}^{m} a_k \chi_{A_k} \leq c + f$.

Let $b_k = \max\{a_k, c\}$ for $k = 1, \dots, m$. Then $b_k - c \geq 0$ for $k = 1, \dots, m$ and

$$\sum_{k=1}^{m} (b_k - c)\chi_{A_k}(x) \leq f(x)$$

for every $x \in X$. The definition of the integral of f implies

$$\sum_{k=1}^{m} (b_k - c)\mu(A_k) \leq \int f\, d\mu,$$

which implies

$$\sum_{k=1}^{m} a_k \mu(A_k) \leq c\mu(X) + \int f\, d\mu.$$

In the inequality above, taking the supremum over all choices satisfying 3.16 shows that

$$\int (c + f)\, d\mu \leq c\mu(X) + \int f\, d\mu,$$

completing the proof. ∎

Monotone Convergence Theorem

The next result allows us to interchange limits and integrals in certain circumstances. We will see more theorems of this nature in the next section.

3.17 *Monotone Convergence Theorem*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $0 \leq f_1 \leq f_2 \leq \cdots$ is an increasing sequence of $\mathcal{S}$-measurable functions. Define $f\colon X \to [0, \infty]$ by

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

Then

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

Idea of this proof: Use Egorov's Theorem to show that the convergence is uniform off a set where the integral is small.

Proof The function f is $\mathcal{S}$-measurable (by 2.50). We will assume that $\int f \, d\mu < \infty$. The case where $\int f \, d\mu = \infty$ uses similar ideas and is left to the reader as an exercise.

Suppose $\varepsilon > 0$. The definition of $\int f \, d\mu$ implies that there exists a simple measurable function $h\colon X \to [0, \infty)$ such that $h \leq f$ and

3.18
$$\int f \, d\mu - \int h \, d\mu < \frac{\varepsilon}{3}.$$

Let H equal the maximum value of the function h (which takes on only finitely many values).

Let $B = \{x \in X : h(x) \neq 0\}$. Then $\mu(B) < \infty$ (because h is a simple function and $\int h \, d\mu \leq \int f \, d\mu < \infty$). Thus by Egorov's Theorem (2.76), there exists a set $E \subset B$ such that $E \in \mathcal{S}$ and $\mu(B \setminus E) < \frac{\varepsilon}{3H}$ and $f_1, f_2, \ldots$ converges uniformly to f on E. Now for each $n \in \mathbf{Z}^+$ we have

$$\begin{aligned}
\int h \, d\mu - \int f_n \, d\mu &= \int \chi_{B \setminus E} h \, d\mu + \int \chi_E h \, d\mu - \int f_n \, d\mu \\
&\leq \int \chi_{B \setminus E} h \, d\mu + \int \chi_E h \, d\mu - \int \chi_E f_n \, d\mu \\
&\leq \frac{\varepsilon}{3} + \int \chi_E h \, d\mu - \int \chi_E f_n \, d\mu, \qquad (3.19)
\end{aligned}$$

where the first line comes from 3.13, the second line comes from 3.10, and the third line comes from 3.11.

If n is sufficiently large, then

$$\begin{aligned}
h(x) &\leq f(x) \\
&< \frac{\varepsilon}{3\mu(E)} + f_n(x)
\end{aligned}$$

for all $x \in E$ (because $f_1, f_2, \ldots$ converges uniformly to f on E). Multiplying the inequality above by χ_E and then integrating gives

$$\int \chi_E h \, d\mu \leq \int \chi_E \Big(\frac{\varepsilon}{3\mu(E)} + f_n(x) \Big) \, d\mu$$
$$\leq \frac{\varepsilon}{3} + \int \chi_E f_n \, d\mu$$

for n sufficiently large, where the last line comes from 3.14 applied to the measure on E that equals μ restricted to the $\mathcal{S}$-measurable subsets of E.

Using the last inequality with 3.19 shows that for n sufficiently large, we have

$$\int h \, d\mu - \int f_n \, d\mu < \frac{2\varepsilon}{3}.$$

Adding the inequality above to 3.18 gives

$$\int f \, d\mu - \int f_n \, d\mu < \varepsilon,$$

for n sufficiently large. Because $f_n \leq f$, we have $\int f_n \, d\mu \leq \int f \, d\mu$. Thus the inequality above completes the proof. ■

3.20 **Example** *Monotone Convergence Theorem fails for decreasing sequence*

Suppose λ is Lebesgue measure on $\mathbf{R}$. Let f_n be the constant function $\frac{1}{n}$ on $\mathbf{R}$, and let f be the constant function 0 on $\mathbf{R}$. Then all these functions are nonnegative, $f_1 \geq f_2 \geq \cdots$, and $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in \mathbf{R}$. However,

$$\lim_{n\to\infty} \int f_n \, d\lambda \neq \int f \, d\lambda$$

because $\int f_n \, d\lambda = \infty$ and $\int f \, d\lambda = 0$.

The result we just proved is traditionally and unfortunately called the Monotone Convergence Theorem. However, it should be called the Increasing Convergence Theorem because it does not hold for decreasing sequences of nonnegative functions, as shown by this example.

Now we can prove that integration is additive on nonnegative functions.

3.21 ***Integral of a sum of nonnegative functions***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \colon X \to [0, \infty]$ are $\mathcal{S}$-measurable functions. Then

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

Proof We already know (see 3.13) that the result is true for simple $\mathcal{S}$-measurable functions. Now approximate f and g by increasing sequences of simple nonnegative $\mathcal{S}$-measurable functions and use the Monotone Convergence Theorem.

More precisely, let $f_1, f_2, \ldots$ and $g_1, g_2, \ldots$ be increasing sequences of simple nonnegative $\mathcal{S}$-measurable functions such that $\lim_{n\to\infty} f_n(x) = f(x)$ and $\lim_{n\to\infty} g_n(x) = g(x)$ (see 2.81). Then $f_1 + g_1, f_2 + g_2, \ldots$ is an increasing sequence of simple nonnegative $\mathcal{S}$-measurable functions such that

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

Thus the Monotone Convergence Theorem (3.17) and 3.13 imply

$$\begin{aligned}\int (f+g)\,d\mu &= \lim_{n\to\infty}\int (f_n+g_n)\,d\mu \\ &= \lim_{n\to\infty}\int f_n\,d\mu + \lim_{n\to\infty}\int g_n\,d\mu \\ &= \int f\,d\mu + \int g\,d\mu,\end{aligned}$$

as desired. ∎

Integration of Real-Valued Functions

The following definition gives us a standard way to write an arbitrary real-valued function as the difference of two nonnegative functions.

3.22 Definition f^+; f^-

Suppose $f\colon X\to[-\infty,\infty]$ is a function. Define functions f^+ and f^- from X to $[0,\infty]$ by

$$f^+(x)=\begin{cases} f(x) & \text{if } f(x)\geq 0,\\ 0 & \text{if } f(x)<0\end{cases} \quad\text{and}\quad f^-(x)=\begin{cases} 0 & \text{if } f(x)\geq 0,\\ -f(x) & \text{if } f(x)<0.\end{cases}$$

Note that if $f\colon X\to[-\infty,\infty]$ is a function, then

$$f=f^+-f^- \quad\text{and}\quad |f|=f^++f^-.$$

The decomposition above allows us to extend our definition of integration to functions that take on negative as well as positive values.

3.23 Definition *integral of a real-valued function*

Suppose $(X,\mathcal{S},\mu)$ is a measure space and $f\colon X\to[-\infty,\infty]$ is a measurable function such that at least one of $\int f^+\,d\mu$ and $\int f^-\,d\mu$ is finite. The *integral* of f with respect to μ, denoted $\int f\,d\mu$, is defined by

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

If $f\geq 0$, then $f^+=f$ and $f^-=0$ and thus this definition is consistent with the previous definition of the integral of a nonnegative function.

3.24 Example *a function whose integral is not defined*

Suppose $f\colon \mathbf{R}\to\mathbf{R}$ is the function defined by

$$f(x)=\begin{cases} 1 & \text{if } x\geq 0,\\ -1 & \text{if } x<0\end{cases}$$

and λ is Lebesgue measure on $\mathbf{R}$. Then $\int f\,d\lambda$ is not defined because $\int f^+\,d\lambda=\infty$ and $\int f^-\,d\lambda=\infty$.

The next result says that the integral of a number times a function is exactly what we expect.

3.25 *Integration is homogeneous*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f : X \to [-\infty, \infty]$ is a function such that $\int f \, d\mu$ is defined. If $c \in \mathbf{R}$, then

$$\int cf \, d\mu = c \int f \, d\mu.$$

Proof Suppose $c \geq 0$. Then

$$\begin{aligned}
\int cf \, d\mu &= \int (cf)^+ \, d\mu - \int (cf)^- \, d\mu \\
&= \int cf^+ \, d\mu - \int cf^- \, d\mu \\
&= c\Bigl(\int f^+ \, d\mu - \int f^- \, d\mu\Bigr) \\
&= c \int f \, d\mu,
\end{aligned}$$

where the third line comes from 3.12.

Now suppose $c < 0$. Then $-c > 0$ and

$$\begin{aligned}
\int cf \, d\mu &= \int (cf)^+ \, d\mu - \int (cf)^- \, d\mu \\
&= \int (-c)f^- \, d\mu - \int (-c)f^+ \, d\mu \\
&= (-c)\Bigl(\int f^- \, d\mu - \int f^+ \, d\mu\Bigr) \\
&= c \int f \, d\mu,
\end{aligned}$$

completing the proof. ■

Now we can prove that integration with respect to a measure has the additive property that is required for a good theory of integration.

3.26 *Integration is additive*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g : X \to [-\infty, \infty]$ are $\mathcal{S}$-measurable functions such that $\int |f| \, d\mu < \infty$ and $\int |g| \, d\mu < \infty$. Then

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

Proof Clearly

$$\begin{aligned}(f+g)^+ - (f+g)^- &= f+g \\ &= f^+ - f^- + g^+ - g^-.\end{aligned}$$

Thus

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+.$$

Both sides of the equation above are sums of nonnegative functions. Thus integrating both sides with respect to μ and using 3.21 gives

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

Rearranging the equation above gives

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

where the left side is not of the form $\infty - \infty$ because $(f+g)^+ \leq f^+ + g^+$ and $(f+g)^- \leq f^- + g^-$. The equation above can be rewritten as

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

completing the proof. ■

The inequality in the next result receives frequent use.

3.27 *Absolute value of integral $\leq$ integral of absolute value*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \colon X \to [-\infty, \infty]$ is a function such that $\int f \, d\mu$ is defined. Then

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

Proof Because $\int f \, d\mu$ is defined, f is an $\mathcal{S}$-measurable function and at least one of $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ is finite. Thus

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

as desired. ■

EXERCISES 3A

1 Suppose X is a set, $\mathcal{S}$ is a σ-algebra on X, and $c \in X$. Define the Dirac measure δ_c on $(X, \mathcal{S})$ by

$$\delta_c(E) = \begin{cases} 1 & \text{if } c \in E, \\ 0 & \text{if } c \notin E. \end{cases}$$

Prove that if $f\colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, then $\int f \, d\delta_c = f(c)$.
[*Careful:* $\{c\}$ *may not be in* $\mathcal{S}$.]

2 Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to [0, \infty]$ is an $\mathcal{S}$-measurable function. Prove that

$$\int f \, d\mu > 0 \text{ if and only if } \mu(\{x \in X : f(x) > 0\}) > 0.$$

3 Give an example of a Borel measurable function $f\colon [0,1] \to (0, \infty)$ such that $L(f, [0,1]) = 0$.
[*Recall that* $L(f, [0,1])$ *denotes the lower Riemann integral, which was defined in Section 1A. If* λ *is Lebesgue measure on* $[0,1]$, *then the previous exercise states that* $\int f \, d\lambda > 0$ *for this function* f, *which is what we expect of a positive function. Thus even though both* $L(f, [0,1])$ *and* $\int f \, d\lambda$ *are defined by taking the supremum of approximations from below, Lebesgue measure captures the right behavior for this function* f *and the lower Riemann integral does not.*]

4 Verify the assertion that integration with respect to counting measure is summation (Example 3.5).

5 Suppose X is a set, $\mathcal{S}$ is the σ-algebra of all subsets of X, and $w\colon X \to [0, \infty]$ is a function. Define a measure μ on $(X, \mathcal{S})$ by

$$\mu(E) = \sum_{x \in E} w(x)$$

for $E \subset X$. Prove that if $f\colon X \to [0, \infty]$ is a function, then

$$\int f \, d\mu = \sum_{x \in X} w(x) f(x),$$

where the infinite sums above are defined as the supremum of all sums over finite subsets of X.

6 Prove the Monotone Convergence Theorem (3.17) in the case when $\int f \, d\mu = \infty$.

7 Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f\colon X \to [0, \infty]$ is an $\mathcal{S}$-measurable function. Define $\mu_f\colon \mathcal{S} \to [0, \infty]$ by

$$\mu_f(A) = \int \chi_A f \, d\mu$$

for $A \in \mathcal{S}$. Prove that μ_f is a measure on $(X, \mathcal{S})$.

8 Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \dots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions. Define $f\colon X \to [0, \infty]$ by $f(x) = \sum_{n=1}^{\infty} f_n(x)$. Prove that

$$\int f \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu.$$

9 Give an example to show that the Monotone Convergence Theorem (3.17) can fail if the hypothesis that $f_1, f_2, \dots$ are nonnegative functions is dropped.

10 Give an example of a sequence $x_1, x_2, \dots$ of real numbers such that

$$\lim_{M \to \infty} \sum_{n=1}^{M} x_n \text{ exists in } \mathbf{R},$$

but $\int x \, d\mu$ is not defined, where μ is counting measure on $\mathbf{Z}^+$ and x is the function from $\mathbf{Z}^+$ to $\mathbf{R}$ defined by $x(n) = x_n$.

For $x_1, x_2, \dots$ a sequence in $[-\infty, \infty]$, define $\liminf_{n \to \infty} x_n$ by

$$\liminf_{n \to \infty} x_n = \lim_{n \to \infty} \inf\{x_n, x_{n+1}, \dots\}.$$

Note that $\inf\{x_n, x_{n+1}, \dots\}$ is an increasing function of n; thus the limit above on the right exists in $[-\infty, \infty]$.

11 Suppose that $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \dots$ is a sequence of nonnegative $\mathcal{S}$-measurable functions on X. Define a function $f\colon X \to [0, \infty]$ by $f(x) = \liminf_{n \to \infty} f_n(x)$. Prove that

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

[*The result above is called Fatou's Lemma. Some textbooks prove Fatou's Lemma and then use it to prove the Monotone Convergence Theorem. Here we are taking the reverse approach—you should be able to use the Monotone Convergence Theorem to give a clean proof of Fatou's Lemma.*]

12 Henri Lebesgue wrote the following about his method of integration:

> I have to pay a certain sum, which I have collected in my pocket. I take the bills and coins out of my pocket and give them to the creditor in the order I find them until I have reached the total sum. This is the Riemann integral. But I can proceed differently. After I have taken all the money out of my pocket I order the bills and coins according to identical values and then I pay the several heaps one after the other to the creditor. This is my integral.

Use 3.8 to explain what Lebesgue meant and to explain why integration of a function with respect to a measure can be thought of as partitioning the range of the function, while Riemann integration depends upon partitioning the domain of the function.
[*The quote above is taken from page 796 of* The Princeton Companion to Mathematics, *edited by Timothy Gowers.*]

3B Limits of Integrals & Integrals of Limits

This section will focus on interchanging limits and integrals. Those tools will allow us to characterize the Riemann integrable functions in terms of Lebesgue measure. We will also develop some approximation tools that will be useful in later chapters.

Bounded Convergence Theorem

We begin this section by introducing some useful notation.

3.28 Definition ***integration on a subset***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $E \in \mathcal{S}$. If $f \colon X \to [-\infty, \infty]$ is an $\mathcal{S}$-measurable function, then $\int_E f \, d\mu$ is defined by

$$\int_E f \, d\mu = \int \chi_E f \, d\mu$$

if the right side of the equation above is defined; otherwise $\int_E f \, d\mu$ is undefined.

Alternatively, you can think of $\int_E f \, d\mu$ as $\int f|_E \, d\mu_E$, where μ_E is the measure obtained by restricting μ to the elements of $\mathcal{S}$ that are contained in E.

Notice that according to the definition above, the notation $\int_X f \, d\mu$ means the same as $\int f \, d\mu$.

The following easy result illustrates the use of this new notation. We now need to adopt the convention that $0 \cdot \infty$ and $\infty \cdot 0$ should both be interpreted to be 0. Notice that both $0 \cdot \infty$ and $\infty \cdot 0$ could appear in the conclusion of the next result.

3.29 ***Bounding an integral***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $E \in \mathcal{S}$, and $f \colon X \to [-\infty, \infty]$ is a function such that $\int_E f \, d\mu$ is defined. Then

$$\left| \int_E f \, d\mu \right| \leq \mu(E) \sup_{x \in E} |f(x)|.$$

Proof Let $c = \sup_{x \in E} |f(x)|$. We have

$$\begin{aligned} \left| \int_E f \, d\mu \right| &= \left| \int \chi_E f \, d\mu \right| \\ &\leq \int \chi_E |f| \, d\mu \\ &\leq \int c \chi_E \, d\mu \\ &= c\mu(E), \end{aligned}$$

where the second line comes from 3.27, the third line comes from 3.10, and the fourth line comes from 3.8. ∎

The next result could be proved as a special case of the Dominated Convergence Theorem (3.35), which we will prove later in this section. Thus you could skip the proof here. However, sometimes you get more insight by seeing an easier proof of an important special case. Thus you may want to read the easy proof of the Bounded Convergence Theorem that is presented next.

3.30 *Bounded Convergence Theorem*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose $f \colon X \to \mathbf{R}$ is $\mathcal{S}$-measurable and $f_1, f_2, \dots$ are $\mathcal{S}$-measurable functions from X to $\mathbf{R}$ such that

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

for all $x \in X$. If there exists $c \in (0, \infty)$ such that

$$|f_n(x)| \leq c$$

for all $n \in \mathbf{Z}^+$ and all $x \in X$, then

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

Note the key role of Egorov's Theorem, which states that pointwise convergence is close to uniform convergence, in the proofs involving interchanging limits and integrals.

Proof Suppose c satisfies the hypothesis of this theorem. Let $\varepsilon > 0$. By Egorov's Theorem (2.76), there exists $E \in \mathcal{S}$ such that $\mu(X \setminus E) < \frac{\varepsilon}{4c}$ and $f_1, f_2, \dots$ converges uniformly to f on E. Now

$$\begin{aligned} \left| \int f_n \, d\mu - \int f \, d\mu \right| &= \left| \int_{X \setminus E} f_n \, d\mu - \int_{X \setminus E} f \, d\mu + \int_E (f_n - f) \, d\mu \right| \\ &\leq \int_{X \setminus E} |f_n| \, d\mu + \int_{X \setminus E} |f| \, d\mu + \int_E |f_n - f| \, d\mu \\ &< \frac{\varepsilon}{2} + \mu(E) \sup_{x \in E} |f_n(x) - f(x)|, \end{aligned}$$

where the last inequality follows from 3.29. Because $f_1, f_2, \dots$ converges uniformly to f on E and $\mu(E) < \infty$, the right side of the inequality above is less than ε for n sufficiently large, which completes the proof. ■

Sets of Measure 0 in Integration Theorems

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. If $f, g \colon X \to [-\infty, \infty]$ are $\mathcal{S}$-measurable functions and

$$\mu(\{x \in X : f(x) \neq g(x)\}) = 0,$$

then the definition of an integral implies that $\int f \, d\mu = \int g \, d\mu$ (or both integrals are undefined). Because what happens on a set of measure 0 often does not matter, the following definition is useful.

3.31 Definition ***almost every***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. A set $E \in \mathcal{S}$ is said to contain *μ-almost every* element of X if $\mu(X \setminus E) = 0$. If the measure μ is clear from the context, then the phrase *almost every* is usually used.

For example, almost every real number is irrational (with respect to the usual Lebesgue measure on $\mathbf{R}$) because $|\mathbf{Q}| = 0$.

Theorems about integrals can almost always be relaxed so that the hypotheses apply only almost everywhere instead of everywhere. For example, consider the Bounded Convergence Theorem (3.30), one of whose hypotheses is that

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

for all $x \in X$. Suppose that the hypotheses of the Bounded Convergence Theorem hold except that the equation above holds only almost everywhere, meaning there is a set $E \in \mathcal{S}$ such that $\mu(X \setminus E) = 0$ and the equation above holds for all $x \in E$. Define new functions $g_1, g_2, \dots$ and g by

$$g_n(x) = \begin{cases} f_n(x) & \text{if } x \in E, \\ 0 & \text{if } x \in X \setminus E \end{cases} \quad \text{and} \quad g(x) = \begin{cases} f(x) & \text{if } x \in E, \\ 0 & \text{if } x \in X \setminus E. \end{cases}$$

Then

$$\lim_{n\to\infty} g_n(x) = g(x)$$

for all $x \in X$. Hence the Bounded Convergence Theorem implies that

$$\lim_{n\to\infty} \int g_n \, d\mu = \int g \, d\mu,$$

which immediately implies that

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

because $\int g_n \, d\mu = \int f_n \, d\mu$ and $\int g \, d\mu = \int f \, d\mu$.

Dominated Convergence Theorem

The next result tells us that if a nonnegative function has a finite integral, then its integral over all small sets (in the sense of measure) is small.

3.32 ***Integrals on small sets are small***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $g \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $\int g \, d\mu < \infty$. Then for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\int_B g \, d\mu < \varepsilon$$

for every set $B \in \mathcal{S}$ such that $\mu(B) < \delta$.

Proof Suppose $\varepsilon > 0$. Let h be a simple $\mathcal{S}$-measurable function such that $0 \leq h \leq g$ and

$$\int g \, d\mu - \int h \, d\mu < \frac{\varepsilon}{2};$$

the existence h with these properties follows from the definition of the integral. Let

$$H = \max\{h(x) : x \in X\}$$

and let $\delta > 0$ be such that $H\delta < \frac{\varepsilon}{2}$.

Suppose $B \in \mathcal{S}$ and $\mu(B) < \delta$. Then

$$\begin{aligned} \int_B g \, d\mu &= \int_B (g - h) \, d\mu + \int_B h \, d\mu \\ &\leq \int (g - h) \, d\mu + H\mu(B) \\ &< \frac{\varepsilon}{2} + H\delta \\ &< \varepsilon, \end{aligned}$$

as desired. ■

Some theorems, such as Egorov's Theorem (2.76) have as a hypothesis that the measure of the entire space is finite. The next result sometimes allows us to get around this hypothesis by restricting attention to a key set of finite measure.

3.33 *Integrable functions live mostly on sets of finite measure*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $g \colon X \to [0, \infty]$ is $\mathcal{S}$-measurable, and $\int g \, d\mu < \infty$. Then for every $\varepsilon > 0$, there exists $E \in \mathcal{S}$ such that $\mu(E) < \infty$ and

$$\int_{X \setminus E} g \, d\mu < \varepsilon.$$

Proof Suppose $\varepsilon > 0$. By the definition of the integral of a nonnegative function (see 3.2), there exists a simple $\mathcal{S}$-measurable function $h \colon X \to [0, \infty)$ such that $h \leq g$ and

3.34
$$\int g \, d\mu - \int h \, d\mu < \varepsilon.$$

Let $E = \{x \in X : h(x) \neq 0\}$. Then $\mu(E) < \infty$ (because otherwise we would have $\int g \, d\mu = \infty$). Because $h(x) = 0$ for all $x \in X \setminus E$, we have

$$\begin{aligned} \int_{X \setminus E} g \, d\mu &= \int_{X \setminus E} (g - h) \, d\mu \\ &\leq \int (g - h) \, d\mu \\ &< \varepsilon, \end{aligned}$$

where the last inequality comes from 3.34. ■

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f_1, f_2, \ldots$ is a sequence of $\mathcal{S}$-measurable functions on X such that $\lim_{n\to\infty} f_n(x) = f(x)$ for all (or almost all) $x \in X$. In general, it is not true that $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$; see Exercises 1 and 2.

We already have two good theorems about interchanging limits and integrals. However, both of these theorems have restrictive hypotheses. Specifically, the Monotone Convergence Theorem (3.17) requires all the functions to be nonnegative and it requires the sequence of functions to be increasing. The Bounded Convergence Theorem (3.30) requires the measure of the whole space to be finite and it requires the sequence of functions to be uniformly bounded by a constant.

The next theorem is the grand result in this area. It does not require the sequence of functions to be nonnegative, it does not require the sequence of functions to be increasing, it does not require the measure of the whole space to be finite, and it does not require the sequence of functions to be uniformly bounded. All these hypotheses are replaced only by a requirement that the sequence of functions is pointwise bounded by a function with a finite integral.

Notice that the Bounded Convergence Theorem follows immediately from the result below (take g to be an appropriate constant function and use the hypothesis in the Bounded Convergence Theorem that $\mu(X) < \infty$).

3.35 *Dominated Convergence Theorem*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space, $f \colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable, and $f_1, f_2, \ldots$ are $\mathcal{S}$-measurable functions from X to $[-\infty, \infty]$ such that

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

for almost every $x \in X$. If there exists an $\mathcal{S}$-measurable function $g \colon X \to [0, \infty]$ such that

$$\int g \, d\mu < \infty \quad \text{and} \quad |f_n(x)| \le g(x)$$

for every $n \in \mathbf{Z}^+$ and almost every $x \in X$, then

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

Proof Suppose $g \colon X \to [0, \infty]$ satisfies the hypotheses of this theorem. If $E \in \mathcal{S}$, then

$$\begin{aligned}\left| \int f_n \, d\mu - \int f \, d\mu \right| &= \left| \int_{X\setminus E} f_n \, d\mu - \int_{X\setminus E} f \, d\mu + \int_E f_n \, d\mu - \int_E f \, d\mu \right| \\ &\le \left| \int_{X\setminus E} f_n \, d\mu \right| + \left| \int_{X\setminus E} f \, d\mu \right| + \left| \int_E f_n \, d\mu - \int_E f \, d\mu \right| \\ &\le 2 \int_{X\setminus E} g \, d\mu + \left| \int_E f_n \, d\mu - \int_E f \, d\mu \right|. \qquad (3.36)\end{aligned}$$

Case 1: Suppose $\mu(X) < \infty$.

Let $\varepsilon > 0$. By 3.32, there exists $\delta > 0$ such that

3.37 $$\int_B g\, d\mu < \frac{\varepsilon}{4}$$

for every set $B \in \mathcal{S}$ such that $\mu(B) < \delta$. By Egorov's Theorem (2.76), there exists a set $E \in \mathcal{S}$ such that $\mu(X \setminus E) < \delta$ and $f_1, f_2, \dots$ converges uniformly to f on E. Now 3.36 and 3.37 imply that

$$\Big|\int f_n\, d\mu - \int f\, d\mu\Big| < \frac{\varepsilon}{2} + \Big|\int_E (f_n - f)\, d\mu\Big|.$$

Because $f_1, f_2, \dots$ converges uniformly to f on E and $\mu(E) < \infty$, the last term on the right is less than $\frac{\varepsilon}{2}$ for all sufficiently large n. Thus $\lim_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu$, completing the proof of case 1.

Case 2: Suppose $\mu(X) = \infty$.

Let $\varepsilon > 0$. By 3.33, there exists $E \in \mathcal{S}$ such that $\mu(E) < \infty$ and

$$\int_{X\setminus E} g\, d\mu < \frac{\varepsilon}{4}.$$

The inequality above and 3.36 imply that

$$\Big|\int f_n\, d\mu - \int f\, d\mu\Big| < \frac{\varepsilon}{2} + \Big|\int_E f_n\, d\mu - \int_E f\, d\mu\Big|.$$

By case 1 as applied to the sequence $f_1|_E, f_2|_E, \dots$, the last term on the right is less than $\frac{\varepsilon}{2}$ for all sufficiently large n. Thus $\lim_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu$, completing the proof of case 2. ∎

Riemann Integrals and Lebesgue Integrals

We can now use the tools we have developed to characterize the Riemann integrable functions. In the theorem below, the left side of the last equation denotes the Riemann integral.

3.38 *Riemann integrable $\Longleftrightarrow$ continuous almost everywhere*

Suppose $a < b$ and $f\colon [a,b] \to \mathbf{R}$ is a bounded function. Then f is Riemann integrable if and only if

$$|\{x \in [a,b] : f \text{ is not continuous at } x\}| = 0.$$

Furthermore, if f is Riemann integrable and λ denotes Lebesgue measure on $\mathbf{R}$, then

$$\int_a^b f = \int_{[a,b]} f\, d\lambda.$$

Proof Suppose $n \in \mathbf{Z}^+$. Consider the partition P_n that divides $[a,b]$ into 2^n subintervals of equal size. Let $I_1, \dots, I_{2^n}$ be the corresponding closed subintervals, each of length $(b-a)/2^n$. Let

3.39 $$g_n = \sum_{k=1}^{2^n} \big(\inf_{x \in I_k} f(x)\big)\chi_{I_k} \quad \text{and} \quad h_n = \sum_{k=1}^{2^n} \big(\sup_{x \in I_k} f(x)\big)\chi_{I_k}.$$

The lower and upper Riemann sums of f for the partition P_n are given by integrals. Specifically,

3.40 $$L(f, P_n, [a,b]) = \int_{[a,b]} g_n \, d\lambda \quad \text{and} \quad U(f, P_n, [a,b]) = \int_{[a,b]} h_n \, d\lambda,$$

where λ is Lebesgue measure on $\mathbf{R}$.

The definitions of g_n and h_n given in 3.39 are actually just a first draft of the definitions. A slight problem arises at each point that is in two of the intervals $I_1, \dots, I_{2^k}$ (in other words, at endpoints of these intervals other than a and b). At each of these points, change the value of g_n to be the infimum of f over the union of the two intervals that contain the point, and change the value of h_n to be the supremum of f over the union of the two intervals that contain the point. This change modifies g_n and h_n on only a finite number of points. Thus the integrals in 3.40 are not affected. This change is needed in order to make 3.42 true (otherwise the two sets in 3.42 might differ by at most countably many points, which would not really change the proof but which would not be as aesthetically pleasing).

Clearly $g_1 \le g_2 \le \cdots$ is an increasing sequence of functions and $h_1 \le h_2 \le \cdots$ is a decreasing sequence of functions on $[a,b]$. Define functions $f^{\mathrm{L}} \colon [a,b] \to \mathbf{R}$ and $f^{\mathrm{U}} \colon [a,b] \to \mathbf{R}$ by

$$f^{\mathrm{L}}(x) = \lim_{n \to \infty} g_n(x) \quad \text{and} \quad f^{\mathrm{U}}(x) = \lim_{n \to \infty} h_n(x).$$

Taking the limit as $n \to \infty$ of both equations in 3.40 (see Exercise 9 in Section 1A) and using the Bounded Convergence Theorem (3.30), we see that

3.41 $$L(f, [a,b]) = \int_{[a,b]} f^{\mathrm{L}} \, d\lambda \quad \text{and} \quad U(f, [a,b]) = \int_{[a,b]} f^{\mathrm{U}} \, d\lambda.$$

Note that $f^{\mathrm{L}} \le f \le f^{\mathrm{U}}$. Now 3.41 implies that f is Riemann integrable if and only if

$$\int_{[a,b]} (f^{\mathrm{U}} - f^{\mathrm{L}}) \, d\lambda = 0,$$

which happens if and only if

$$|\{x \in [a,b] : f^{\mathrm{U}}(x) \ne f^{\mathrm{L}}(x)\}| = 0.$$

However,

3.42 $$\{x \in [a,b] : f^{\mathrm{U}}(x) \ne f^{\mathrm{L}}(x)\} = \{x \in [a,b] : f \text{ is not continuous at } x\},$$

as you should verify, which completes the proof. ■

Approximation by Nice Functions

In the next definition, the notation $\|f\|_1$ should be $\|f\|_{1,\mu}$ because it depends upon the measure μ as well as upon f. However, μ is usually clear from the context.

In some books, you may see the notation $\mathcal{L}^1(X, \mathcal{S}, \mu)$ instead of $\mathcal{L}^1(\mu)$. Here we will use the shorter notation because μ determines $\mathcal{S}$ (as the domain of μ) and X (as the unique largest set in the σ-algebra $\mathcal{S}$).

3.43 Definition $\|f\|_1$; $\mathcal{L}^1(\mu)$

Suppose $(X, \mathcal{S}, \mu)$ is a measure space. If $f\colon X \to [-\infty, \infty]$ is $\mathcal{S}$-measurable, then the $\mathcal{L}^1$-*norm* of f is denoted by $\|f\|_1$ and is defined by

$$\|f\|_1 = \int |f| \, d\mu.$$

The *Lebesgue space* $\mathcal{L}^1(\mu)$ is defined by

$$\mathcal{L}^1(\mu) = \{f : f \text{ is an } \mathcal{S}\text{-measurable function from } X \text{ to } \mathbf{R} \text{ and } \|f\|_1 < \infty.\}$$

3.44 Example $\mathcal{L}^1(\mu)$ *functions that take on only finitely many values*

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $E_1, \dots, E_n$ are disjoint subsets of X. Suppose $a_1, \dots, a_n$ are nonzero real numbers. Then

$$a_1 \chi_{E_1} + \cdots + a_n \chi_{E_n} \in \mathcal{L}^1(\mu)$$

if and only if $E_k \in \mathcal{S}$ and $\mu(E_k) < \infty$ for all $k \in \{1, \dots, n\}$, as you should verify.

3.45 Example ℓ^1 *equals* $\mathcal{L}^1$(*counting measure on* $\mathbf{Z}^+$)

If μ equals counting measure on $\mathbf{Z}^+$ and $x = x_1, x_2, \dots$ is a sequence of real numbers (thought of as a function on $\mathbf{Z}^+$), then $\|x\|_1 = \sum_{n=1}^{\infty} |x_n|$. Is this case, $\mathcal{L}^1(\mu)$ is often denoted by ℓ^1 (pronounced "little-el-one"). In other words, ℓ^1 is the set of all sequences $x_1, x_2, \dots$ of real numbers such that $\sum_{n=1}^{\infty} |x_n| < \infty$.

The easy proof of the following result is left to the reader.

3.46 ***Properties of the*** $\mathcal{L}^1$***-norm***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f, g \in \mathcal{L}^1(\mu)$. Then

- $\|f\|_1 \geq 0$;
- $\|f\|_1 = 0$ if and only $f(x) = 0$ for almost every $x \in X$;
- $\|cf\|_1 = |c| \|f\|_1$ for all $c \in \mathbf{R}$;
- $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.

The next result states every function in $\mathcal{L}^1(\mu)$ can be approximated in $\mathcal{L}^1$-norm by measurable functions that take on only finitely many values.

3.47 ***Approximation by simple functions***

Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \in \mathcal{L}^1(\mu)$. Then for every $\varepsilon > 0$, there exists a simple function $g \in \mathcal{L}^1(\mu)$ such that

$$\|f - g\|_1 < \varepsilon.$$

Proof Suppose $\varepsilon > 0$. The definition of the integral of a nonnegative function implies that there exist simple functions $g_+, g_- \in \mathcal{L}^1(\mu)$ such that $0 \leq g_+ \leq f^+$ and $0 \leq g_- \leq f^-$ and

$$\int (f^+ - g_+)\,d\mu < \frac{\varepsilon}{2} \quad \text{and} \quad \int (f^- - g_-)\,d\mu < \frac{\varepsilon}{2}.$$

Let $g = g_+ - g_-$. Then g is a simple function in $\mathcal{L}^1(\mu)$ and

$$\begin{aligned}\|f - g\|_1 &= \|(f^+ - g_+) - (f^- - g_-)\|_1\\ &= \int (f^+ - g_+)\,d\mu + \int (f^- - g_-)\,d\mu\\ &< \varepsilon,\end{aligned}$$

as desired. ■

3.48 Definition $\mathcal{L}^1(\mathbf{R})$; $\|f\|_1$

- The notation $\mathcal{L}^1(\mathbf{R})$ denotes $\mathcal{L}^1(\lambda)$, where λ is Lebesgue measure on the Borel subsets of $\mathbf{R}$.
- When working with $\mathcal{L}^1(\mathbf{R})$, the notation $\|f\|_1$ denotes the integral of the absolute value of f with respect to Lebesgue measure on $\mathbf{R}$.

3.49 Definition ***step function***

A *step function* is a function $g\colon \mathbf{R} \to \mathbf{R}$ of the form

$$g = a_1\chi_{I_1} + \cdots + a_n\chi_{I_n},$$

where $I_1, \ldots, I_n$ are intervals of $\mathbf{R}$ and $a_1, \ldots, a_n$ are nonzero real numbers.

Suppose g is a step function of the form above and the intervals $I_1, \ldots, I_n$ are disjoint. Then

$$\|g\|_1 = |a_1|\,|I_1| + \cdots + |a_n|\,|I_n|.$$

In particular, $g \in \mathcal{L}^1(\mathbf{R})$ if and only if all the intervals $I_1, \ldots, I_n$ are bounded.

Even though the coefficients $a_1, \dots, a_n$ in the definition of a step function are required to be nonzero, the function 0 that is identically 0 on **R** *is a step function. To see this, take $n = 1$, $a_1 = 1$, and $I_1 = \emptyset$.*

The intervals in the definition of a step function can be open intervals, closed intervals, or half-open intervals. We will be using step functions in integrals, where the inclusion or exclusion of the endpoints of the intervals does not matter.

3.50 *Approximation by step functions*

Suppose $f \in \mathcal{L}^1(\mathbf{R})$. Then for every $\varepsilon > 0$, there exists a step function $g \in \mathcal{L}^1(\mathbf{R})$ such that

$$\|f - g\|_1 < \varepsilon.$$

Proof Suppose $\varepsilon > 0$. By 3.47, there exist Borel subsets $B_1, \dots, B_N$ of **R** and nonzero numbers $a_1, \dots, a_N$ such that $|B_k| < \infty$ for all $k \in \{1, \dots, N\}$ and

$$\Big\|f - \sum_{k=1}^{N} a_k \chi_{B_k}\Big\|_1 < \frac{\varepsilon}{2}.$$

For each $k \in \{1, \dots, N\}$, there is a countable union of disjoint bounded open intervals that contains B_k and whose Lebesgue measure is as close as we want to $|B_k|$ (by the definition of outer measure). For each of those countable unions of disjoint open intervals, there is a finite union of a subcollection of the open intervals whose Lebesgue measure is as close as we want to the Lebesgue measure of the countable union. Thus for each k, there is a set E_k that is a finite union of disjoint bounded intervals such that

$$|E_k \setminus B_k| + |B_k \setminus E_k| < \frac{\varepsilon}{2|a_k|N};$$

in other words,

$$\|\chi_{B_k} - \chi_{E_k}\|_1 < \frac{\varepsilon}{2|a_k|N}.$$

Now

$$\begin{aligned}\Big\|f - \sum_{k=1}^{N} a_k \chi_{E_k}\Big\|_1 &\le \Big\|f - \sum_{k=1}^{N} a_k \chi_{B_k}\Big\|_1 + \Big\|\sum_{k=1}^{N} a_k \chi_{B_k} - \sum_{k=1}^{N} a_k \chi_{E_k}\Big\|_1 \\ &< \frac{\varepsilon}{2} + \sum_{x=1}^{N} |a_k| \|\chi_{B_k} - \chi_{E_k}\|_1 \\ &< \varepsilon,\end{aligned}$$

completing the proof. ■

Luzin's Theorem (2.82 and 2.84) gives a spectacular way to approximate a Borel measurable function by a continuous function. However, the following approximation theorem is usually more useful than Luzin's Theorem. For example, the next result plays a major role in the proof of the Lebesgue Differentiation Theorem (4.11).

3.51 *Approximation by continuous functions*

Suppose $f \in \mathcal{L}^1(\mathbf{R})$. Then for every $\varepsilon > 0$, there exists a continuous function $g\colon \mathbf{R} \to \mathbf{R}$ such that

$$\|f - g\|_1 < \varepsilon$$

and $\{x \in \mathbf{R} : g(x) \neq 0\}$ is a bounded set.

Proof By 3.50, we need only prove this result in the case when f is the characteristic function of a bounded interval. Suppose $f = \chi_{[a,b]}$. Let δ be a number such that $0 < \delta < \varepsilon$. Define $g\colon \mathbf{R} \to \mathbf{R}$ by

$$g(x) = \begin{cases} 0 & \text{if } x \leq a - \delta, \\ \frac{x-(a-\delta)}{\delta} & \text{if } a - \delta < x < a, \\ 1 & \text{if } a \leq x \leq b, \\ \frac{(b+\delta)-x}{\delta} & \text{if } b < x < b + \delta, \\ 0 & \text{if } x \geq b + \delta. \end{cases}$$

Then g is a continuous function, $\{x \in \mathbf{R} : g(x) \neq 0\}$ is a bounded interval, and $\|f - g\|_1 = \delta < \varepsilon$. ■

EXERCISES 3B

1 Give an example of a sequence $f_1, f_2, \ldots$ of functions from $\mathbf{Z}^+$ to $[0, \infty)$ such that

$$\lim_{n\to\infty} f_n(m) = 0$$

for every $m \in \mathbf{Z}^+$ but $\displaystyle\lim_{n\to\infty} \int f_n \, d\mu = 1$, where μ is counting measure on $\mathbf{Z}^+$.

2 Give an example of a sequence $f_1, f_2, \ldots$ of continuous functions from $\mathbf{R}$ to $[0, 1]$ such that

$$\lim_{n\to\infty} f_n(x) = 0$$

for every $x \in \mathbf{R}$ but $\displaystyle\lim_{n\to\infty} \int f_n \, d\lambda = \infty$, where λ is Lebesgue measure on $\mathbf{R}$.

3 Suppose λ is Lebesgue measure on $\mathbf{R}$ and $f\colon \mathbf{R} \to \mathbf{R}$ is a Borel measurable function such that $\int |f| \, d\lambda < \infty$. Define $g\colon \mathbf{R} \to \mathbf{R}$ by

$$g(x) = \int_{(-\infty, x)} f \, d\lambda.$$

Prove that g is uniformly continuous on $\mathbf{R}$.

4 Suppose $a < b$ and $f\colon [a,b] \to \mathbf{R}$ is a bounded Borel measurable function. Let λ denote Lebesgue measure on $\mathbf{R}$. Prove that

$$\int_{[a,b]} f\, d\lambda$$
$$= \inf\{\int_{[a,b]} h\, d\lambda : h \text{ is a simple Borel measurable function and } f \le h\}.$$

5 Let λ denote Lebesgue measure on $\mathbf{R}$. Suppose $f\colon \mathbf{R} \to \mathbf{R}$ is a Borel measurable function such that $\int |f|\, d\lambda < \infty$. Prove that

$$\lim_{n\to\infty} \int_{[-n,n]} f\, d\lambda = \int f\, d\lambda.$$

6 Let λ denote Lebesgue measure on $\mathbf{R}$. Give an example of a continuous function $f\colon [0,\infty) \to \mathbf{R}$ such that $\lim_{t\to\infty} \int_{[0,t]} f\, d\lambda$ exists (in $\mathbf{R}$) but $\int_{[0,\infty)} f\, d\lambda$ is not defined.

7 Let λ denote Lebesgue measure on $\mathbf{R}$. Give an example of a continuous function $f\colon (0,1) \to \mathbf{R}$ such that $\lim_{n\to\infty} \int_{(\frac{1}{n},1)} f\, d\lambda$ exists (in $\mathbf{R}$) but $\int_{(0,1)} f\, d\lambda$ is not defined.

8 Verify the assertion in 3.42.

9 Verify the assertion in Example 3.44.

10 Suppose $(X, \mathcal{S}, \mu)$ is a measure space such that $\mu(X) < \infty$. Suppose p, r are positive numbers with $p < r$. Prove that if $f\colon X \to [0,\infty)$ is an $\mathcal{S}$-measurable function such that $\int f^r\, d\mu < \infty$, then $\int f^p\, d\mu < \infty$.

11 Give an example to show that the result in the previous exercise can be false without the hypothesis that $\mu(X) < \infty$.

12 Suppose $(X, \mathcal{S}, \mu)$ is a measure space and $f \in \mathcal{L}^1(\mu)$. Prove that

$$\{x \in X : f(x) \neq 0\}$$

is the countable union of sets with finite μ-measure.

13 For $f\colon \mathbf{R} \to \mathbf{R}$ and $t \in \mathbf{R}$, define $f_t\colon \mathbf{R} \to \mathbf{R}$ by $f_t(x) = f(x-t)$ [thus if $t > 0$, then the graph of f_t is obtained by shifting the graph of f to the right by t units]. Prove that if $\mathcal{L}^1(\mathbf{R})$, then

$$\lim_{t\to 0} \|f_t - f\|_1 = 0.$$