V. Lebesgue Measure and Abstract Measure Theory, 267-333

DOI: 10.3792/euclid/9781429799997-5

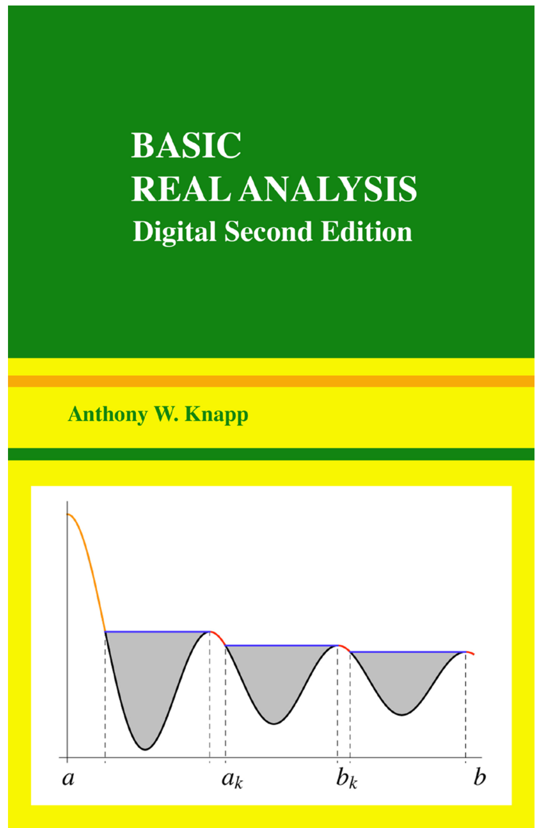
from

Basic Real Analysis

Digital Second Edition

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799997
ISBN: 978-1-4297-9999-7



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/∼aknapp

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"
Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28–01, 26–01, 42–01, 54–01, 34–01, 30–01, 32–01.

First Edition, ISBN-13 978-0-8176-3250-2

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the author

CHAPTER V

Lebesgue Measure and Abstract Measure Theory

Abstract. This chapter develops the basic theory of measure and integration, including Lebesgue measure and Lebesgue integration for the line.

Section 1 introduces measures, including 1-dimensional Lebesgue measure as the primary example, and develops simple properties of them. Sections 2–4 introduce measurable functions and the Lebesgue integral and go on to establish some easy properties of integration and the fundamental theorems about how Lebesgue integration behaves under limit operations.

Sections 5–6 concern the Extension Theorem announced in Section 1 and used as the final step in the construction of Lebesgue measure. The theorem allows σ-finite measures to be extended from algebras of sets to σ-algebras. The theorem is proved in Section 5, and the completion of a measure space is defined in Section 6 and related to the proof of the Extension Theorem.

Section 7 treats Fubini's Theorem, which allows interchange of order of integration under rather general circumstances. This is a deep result. As part of the proof, product measure is constructed and important measurability conditions are established. This section mentions that Fubini's Theorem will be applicable to higher-dimensional Lebesgue measure, but the details are deferred to Chapter VI.

Section 8 extends Lebesgue integration to complex-valued functions and to functions with values in finite-dimensional vector spaces.

Section 9 gives a careful definition of the spaces L^1, L^2, and L^∞ for any measure space, introduces the notion of a normed linear space, and verifies that these three spaces are examples. The main theorem of the section about L^1, L^2, and L^∞ is the completeness of these three spaces as metric spaces. In addition, the section proves a version of Alaoglu's Theorem concerning weak-star convergence.

1. Measures and Examples

In the theory of the Riemann integral, as discussed in Chapter I for $\mathbb{R}^1$ and in Chapter III for $\mathbb{R}^n$, we saw that Riemann integration is a powerful tool when applied to continuous functions. Riemann integration makes sense also when applied to certain kinds of discontinuous functions, but then the theory has some weaknesses.

Without any change in the definitions, one of these is that the theory applies only to bounded functions. Thus we can compute $\int_0^1 x^p\,dx = \big[x^{p+1}/(p+1)\big]_0^1 = (p+1)^{-1}$ for $p \geq 0$, but only the right side makes sense for $-1 < p < 0$. More seriously we made calculations with trigonometric series in Section I.10 and found that $\frac{1}{2}\log\big(\frac{1}{2-2\cos\theta}\big) = \sum_{n=1}^\infty \frac{\cos n\theta}{n}$ and $\frac{1}{2}(\pi-\theta) = \sum_{n=1}^\infty \frac{\sin n\theta}{n}$ for $0 < \theta < 2\pi$.

When we tried to explain these similar-looking identities with Fourier series, we were able to handle the second one because $\frac{1}{2}(\pi - \theta)$ is a bounded function, but we were not able to handle the first one because $\frac{1}{2}\log\left(\frac{1}{2-2\cos\theta}\right)$ is unbounded.

Other weaknesses appeared in Chapters I–IV at certain times: when we always had to arrange for the set of integration to be bounded, when we had no clue which sequences $\{c_n\}$ of Fourier coefficients occurred in the beautiful formula given by Parseval's Theorem, when Fubini's Theorem turned out to be awkward to apply to discontinuous functions, and when the change-of-variables formula did not immediately yield the desired identities even in simple cases like the change from Cartesian coordinates to polar coordinates.

The Lebesgue integral will solve all these difficulties when formed with respect to "Lebesgue measure" in the setting of $\mathbb{R}^n$. In addition, the Lebesgue integral will be meaningful in other settings. For example, the Lebesgue integral will be meaningful on the unit sphere in Euclidean space, while the Riemann integral would always require a choice of coordinates. The Lebesgue integral will be meaningful also in other situations where we can take advantage of some action by a group (such as a rotation group) that is difficult to handle when the setting has to be Euclidean. And the Lebesgue integral will enable us to provide a rigorous foundation for the theory of probability.

There are five ingredients in Lebesgue integration, and these will be introduced in Sections 1–3 of this chapter:

(i) an underlying nonempty set, such as $\mathbb{R}^1$ in the case of Lebesgue integration on the line,
(ii) a distinguished class of subsets, called the "measurable sets," which will form a "σ-ring" or a "σ-algebra,"
(iii) a measure, which attaches a member of $[0, +\infty]$ to each measurable set and which will be "length" in the case of Lebesgue measure on the line,
(iv) the "measurable functions," those functions with values in $\mathbb{R}$ (or some more general space) that we try to integrate,
(v) the integral of a measurable function over a measurable set.

Let us write X for the underlying nonempty set. The important thing about whatever sets are measurable will be that certain simple set-theoretic operations lead from measurable sets to measurable sets. The two main definitions are those of an "algebra" of sets and a "σ-algebra," but we shall refer also to the notions of a "ring" of sets and a "σ-ring" in order to simplify certain technical problems in constructing measures. An **algebra of sets** $\mathcal{A}$ is a set of subsets of X containing $\varnothing$ and X and closed under the operation of forming the union $E \cup F$ of two sets and under taking the complement E^c of a set. An algebra is necessarily closed under intersection $E \cap F$ and difference $E - F = E \cap F^c$. Another operation under which $\mathcal{A}$ is closed is **symmetric difference**, which is defined by

$E \,\Delta\, F = (E - F) \cup (F - E)$; we shall make extensive use of this operation[1] in Section 6 of this chapter.

In practice, despite the effort often needed to define an interesting measure on the sets in an algebra, the closure properties[2] of the algebra are insufficient to deal with questions about limits. For this reason one defines a **σ-algebra** of subsets of X to be an algebra that is closed under countable unions (and hence also countable intersections). Typically a general foundational theorem (Theorem 5.5 below) is used to extend the constructed would-be measure from an algebra to a σ-algebra.

A **ring** $\mathcal{R}$ of subsets of X is a set of subsets closed under finite unions and under difference. Then $\mathcal{R}$ is closed also under the operations of finite intersections, difference, and symmetric difference.[3] A **σ-ring** of subsets of X is a ring of subsets that is closed under countable unions.

EXAMPLES.

(1) $\mathcal{A} = \{\varnothing, X\}$. This is a σ-algebra.

(2) All subsets of X. This is a σ-algebra.

(3) All finite subsets of X. This is a ring. If the complements of such sets are included, the result is an algebra.

(4) All finite and countably infinite subsets of X. This is a σ-ring. If the complements of such sets are included, the result is a σ-algebra.

(5) All **elementary sets** of $\mathbb{R}$. These are all finite disjoint unions of bounded intervals in $\mathbb{R}$ with or without endpoints. This collection is a ring. To see the closure properties, we first verify that any finite union of bounded intervals is a finite disjoint union; in fact, if $I_1, \dots, I_n$ are bounded intervals such that none contains any of the others, then $I_k - \bigcup_{m=1}^{k-1} I_m$ is an interval, and these intervals are disjoint as k varies; also these intervals have the same union as $I_1, \dots, I_n$. Now let $E = \bigcup_i I_i$ and $F = \bigcup_j J_j$ be given. Since $I_i \cap J_j$ is an interval, the identity $E \cap F = \bigcup_{i,j} (I_i \cap J_j)$ shows that $E \cap F$ is a finite union of intervals. Since each $I_i - J_j$ is an interval or the union of two intervals, the identity $E - F = \bigcup_i \bigcap_j (I_i - J_j)$ then shows that $E - F$ is a finite union of intervals.

(6) If $\mathcal{C}$ is an arbitrary class of subsets of X, then there is a unique smallest algebra $\mathcal{A}$ of subsets of X containing $\mathcal{C}$. Similar statements apply to σ-algebras,

[1] For some properties of symmetric difference, see Problem 1 at the end of the chapter.

[2] An algebra of sets really is an algebra in the sense of the discussion of algebras with the Stone–Weierstrass Theorem (Theorem 2.58). The scalars replacing $\mathbb{R}$ or $\mathbb{C}$ are the members of the two-element field $\{0, 1\}$, addition is given by symmetric difference, and multiplication is given by intersection. The additive identity is $\varnothing$, the multiplicative identity is X, and every element is its own negative. Multiplication is commutative.

[3] A ring of sets really is a ring in the sense of modern algebra; addition is given by symmetric difference, and multiplication is given by intersection.

rings, and σ-rings. In fact, consider all algebras of subsets of X containing $\mathcal{C}$. Example 2 shows that there is one. Let $\mathcal{A}$ be the intersection of all these algebras, i.e., the set of all subsets that occur in each of these algebras. If two sets occur in $\mathcal{A}$, they occur in each such algebra, and their intersection is in each algebra. Hence their intersection is in $\mathcal{A}$. Similarly $\mathcal{A}$ is closed under differences.

If $\mathcal{R}$ is a ring of subsets of X, a **set function** is a function $\rho : \mathcal{R} \to \mathbb{R}^*$, where $\mathbb{R}^*$ denotes the extended real-number system as in Section I.1. The set function is **nonnegative** if its values are all in $[0, +\infty]$, it is **additive** if $\rho(\varnothing) = 0$ and if $\rho(E \cup F) = \rho(E) + \rho(F)$ whenever E and F are disjoint sets in $\mathcal{R}$, and it is **completely additive** or **countably additive** if $\rho(\varnothing) = 0$ and if $\rho\big(\bigcup_{n=1}^\infty E_n\big) = \sum_{n=1}^\infty \rho(E_n)$ whenever the sets E_n are pairwise disjoint members of $\mathcal{R}$ with $\bigcup_{n=1}^\infty E_n$ in $\mathcal{R}$. In the definitions of "additive" and "completely additive," it is taken as part of the definition that the sums in question are to be well defined in $\mathbb{R}^*$. Observe that completely additive implies additive, since $\rho(\varnothing) = 0$.

Proposition 5.1. An additive set function ρ on a ring $\mathcal{R}$ of sets has the following properties:

(a) $\rho\big(\bigcup_{n=1}^N E_n\big) = \sum_{n=1}^N \rho(E_n)$ if the sets E_n are pairwise disjoint and are in $\mathcal{R}$.
(b) $\rho(E \cup F) + \rho(E \cap F) = \rho(E) + \rho(F)$ if E and F are in $\mathcal{R}$.
(c) If E and F are in $\mathcal{R}$ and $|\rho(E)| < +\infty$, then $|\rho(E \cap F)| < +\infty$.
(d) If E and F are in $\mathcal{R}$ and if $|\rho(E \cap F)| < +\infty$, then $\rho(E - F) = \rho(E) - \rho(E \cap F)$.
(e) If ρ is nonnegative and if E and F are in $\mathcal{R}$ with $E \subseteq F$, then $\rho(E) \le \rho(F)$.
(f) If ρ is nonnegative and if $E, E_1, \dots, E_N$ are sets in $\mathcal{R}$ such that $E \subseteq \bigcup_{n=1}^N E_n$, then $\rho(E) \le \sum_{n=1}^N \rho(E_n)$.
(g) If ρ is nonnegative and completely additive and if $E, E_1, E_2, \dots$ are sets in $\mathcal{R}$ such that $E \subseteq \bigcup_{n=1}^\infty E_n$, then $\rho(E) \le \sum_{n=1}^\infty \rho(E_n)$.

PROOF. Part (a) follows by induction from the definition. In (b), we have $E \cup F = (E - F) \cup (E \cap F) \cup (F - E)$ disjointly. Application of (a) gives $\rho(E \cup F) = \rho(E - F) + \rho(E \cap F) + \rho(F - E)$, with $+\infty$ and $-\infty$ not both occurring. Adding $\rho(E \cap F)$ to both sides, regrouping terms, and taking into account that $\rho(E) = \rho(E - F) + \rho(E \cap F)$ and $\rho(F) = \rho(F - E) + \rho(E \cap F)$, we obtain (b). The right side of the identity $\rho(E) = \rho(E \cap F) + \rho(E - F)$ cannot be well defined if $\rho(E)$ is finite and $\rho(E \cap F)$ is infinite, and thus (c) follows. In the identity $\rho(E) = \rho(E \cap F) + \rho(E - F)$, we can subtract $\rho(E \cap F)$ from both sides and obtain (d) if $\rho(E \cap F)$ is finite. For (e), the inclusion $E \subseteq F$ forces $F = (F - E) \cup E$ disjointly; then $\rho(F) = \rho(F - E) + \rho(E)$, and (e) follows. In (f), put $F_n = E_n - \bigcup_{k=1}^{n-1} E_k$. Then $E = \bigcup_{n=1}^N (E \cap F_n)$ disjointly, and (a) and

(e) give $\rho(E) = \sum_{n=1}^N \rho(E \cap F_n) \le \sum_{n=1}^N \rho(F_n) \le \sum_{n=1}^N \rho(E_n)$. Conclusion (g) is proved in the same way as (f). $\square$

Proposition 5.2. Let ρ be an additive set function on a ring $\mathcal{R}$ of sets. If ρ is completely additive, then $\rho(E) = \lim \rho(E_n)$ whenever $\{E_n\}$ is an increasing sequence of members of $\mathcal{R}$ with union E in $\mathcal{R}$. Conversely if $\rho(E) = \lim \rho(E_n)$ for all such sequences, then ρ is completely additive.

PROOF. First we prove the direct part of the proposition. For E and E_n as in the statement, let $F_1 = E_1$ and $F_n = E_n - E_{n-1}$ for $n \ge 2$. Then $E_n = \bigcup_{k=1}^n F_k$ disjointly, and $\rho(E_n) = \sum_{k=1}^n \rho(F_k)$ by additivity. Also, $E = \bigcup_{k=1}^\infty F_k$, and complete additivity gives $\rho(E) = \sum_{k=1}^\infty \rho(F_k) = \lim \sum_{k=1}^n \rho(F_k) = \lim \rho(E_n)$. The direct part of the proposition follows.

For the converse let $\{F_n\}$ be a disjoint sequence in $\mathcal{R}$ with union F in $\mathcal{R}$. Put $E_n = \bigcup_{k=1}^n F_k$. Then E_n is an increasing sequence of sets in $\mathcal{R}$ with union F in $\mathcal{R}$. We are given that $\rho(F) = \lim \rho(E_n)$, and we have $\rho(E_n) = \sum_{k=1}^n \rho(F_k)$ by additivity and Proposition 5.1a. Therefore $\rho(F) = \sum_{k=1}^\infty \rho(F_k)$, and we conclude that ρ is completely additive. $\square$

Corollary 5.3. Let ρ be an additive set function on an algebra $\mathcal{A}$ of subsets of X such that $|\rho(X)| < +\infty$. If ρ is completely additive, then $\rho(E) = \lim \rho(E_n)$ whenever $\{E_n\}$ is a decreasing sequence of members of $\mathcal{A}$ with intersection E in $\mathcal{A}$. Conversely if $\lim \rho(E_n) = 0$ whenever $\{E_n\}$ is a decreasing sequence of members of $\mathcal{A}$ with intersection empty, then ρ is completely additive.

PROOF. This follows from Proposition 5.2 by taking complements. $\square$

A **measure** is a nonnegative completely additive set function on a σ-ring of subsets of X. If no ambiguity is possible about the σ-ring, we may refer to a "measure on X." When we use measures to work with integrals, the σ-ring will be taken to be a σ-algebra; if integration were to be defined relative to a σ-ring that is not a σ-algebra, then constant functions would not be measurable.

The assumption that our σ-ring is a σ-algebra for doing integration is no loss of generality. Even when the σ-ring is not a σ-algebra, there is a canonical way of extending a measure from a σ-ring to the smallest σ-algebra containing the σ-ring. Proposition 5.37 at the end of Section 5 gives the details.

EXAMPLES.

(1) For $\{\varnothing, X\}$, define $\mu(X) = a \ge 0$. This is a measure.

(2) For X equal to a countable set and with all subsets in the σ-algebra, attach a weight ≥ 0 to each member of X. Define $\mu(E)$ to be the sum of the weights for the members of E. This is a measure.

(3) For X arbitrary but nonempty, let $\mu(E)$ be the number of points in E, a nonnegative integer or $+\infty$. We refer to μ as **counting measure**.

(4) **Lebesgue measure** m on the ring $\mathcal{R}$ of elementary sets of $\mathbb{R}$. If E is a finite disjoint union of bounded intervals, we let $m(E)$ be the sum of the lengths of the intervals. We need to see that this definition is unambiguous. Consider the special case that $J = I_1 \cup \cdots \cup I_r$ disjointly with I_k extending from a_k to b_k, with or without endpoints. Then we can arrange the intervals in order so that $b_k = a_{k+1}$ for $k = 1, \dots, r-1$. In this case, $m(J) = b_r - a_1$ and $\sum_{k=1}^r m(I_k) = \sum_{k=1}^r (b_k - a_k) = b_r - a_1$. Thus the definition is unambiguous in this special case. If $E = I_1 \cup \cdots \cup I_r = J_1 \cup \cdots \cup J_s$, then the special case gives $m(J_k) = \sum_{j=1}^r m(I_j \cap J_k)$ and hence $\sum_{k=1}^s m(J_k) = \sum_{j,k} m(I_j \cap J_k)$. Reversing the roles of the I_j's and the J_k's, we obtain $\sum_{j=1}^r m(I_j) = \sum_{j,k} m(I_j \cap J_k)$. Thus $\sum_{k=1}^s m(J_k) = \sum_{j=1}^r m(I_j)$, and the definition of m on $\mathcal{R}$ is unambiguous. It is evident that m is nonnegative and additive. We shall prove that m is completely additive on $\mathcal{R}$. Even so, m will not yet be a measure, since $\mathcal{R}$ is not a σ-ring. That step will have to be carried out separately. Proving that m is completely additive on the ring $\mathcal{R}$ uses the fact that m is **regular** on $\mathcal{R}$ in the sense that if E is in $\mathcal{R}$ and if $\epsilon > 0$ is given, then there exist a compact set K in $\mathcal{R}$ and an open set U in $\mathcal{R}$ such that $K \subseteq E \subseteq U$, $m(K) \geq m(E) - \epsilon$, and $m(U) \leq m(E) + \epsilon$: In the special case that E is a single bounded interval with endpoints a and b, we can prove regularity by taking $U = (a - \epsilon/2, b + \epsilon/2)$ and by letting $K = \varnothing$ if $b - a \leq \epsilon$ or $K = [a + \epsilon/2, b - \epsilon/2]$ if $b - a > \epsilon$. In the general case that E is the union of n bounded intervals I_j, choose K_j and U_j for I_j and for the number ϵ/n, and put $K = \bigcup_{j=1}^n K_j$ and $U = \bigcup_{j=1}^n U_j$. Then $m(K) = \sum_{j=1}^n m(K_j) \geq \sum_{j=1}^n \big(m(I_j) - \epsilon/n\big) = m(E) - \epsilon$, and Proposition 5.1f gives $m(U) \leq \sum_{j=1}^n m(U_j) \leq \sum_{j=1}^n \big(m(I_j) + \epsilon/n\big) = m(E) + \epsilon$.

Proposition 5.4. Lebesgue measure m is completely additive on the ring $\mathcal{R}$ of elementary sets in $\mathbb{R}^1$.

PROOF. Let $\{E_n\}$ be a disjoint sequence in $\mathcal{R}$ with union E in $\mathcal{R}$. Since m is nonnegative and additive, Proposition 5.1 gives $m(E) \geq m\big(\bigcup_{k=1}^n E_k\big) = \sum_{k=1}^n m(E_k)$ for every n. Passing to the limit, we obtain $m(E) \geq \sum_{k=1}^\infty m(E_k)$. For the reverse inequality, let $\epsilon > 0$ be given. Choose by regularity a compact member K of $\mathcal{R}$ and open members U_n of $\mathcal{R}$ such that $K \subseteq E$, $U_n \supseteq E_n$ for all n, $m(K) \geq m(E) - \epsilon$, and $m(U_n) \leq m(E_n) + \epsilon/2^n$. Then $K \subseteq \bigcup_{n=1}^\infty U_n$, and the compactness implies that $K \subseteq \bigcup_{n=1}^N U_n$ for some N. Hence $m(E) - \epsilon \leq m(K) \leq \sum_{n=1}^N m(U_n) \leq \sum_{n=1}^N (m(E_n) + \epsilon/2^n) \leq \sum_{n=1}^\infty m(E_n) + \epsilon$. Since ϵ is arbitrary, $m(E) \leq \sum_{n=1}^\infty m(E_n)$, and the proposition follows. $\square$

The smallest σ-ring containing the ring $\mathcal{R}$ of elementary sets in $\mathbb{R}^1$ is in fact a σ-algebra, since $\mathbb{R}^1$ is the countable union of bounded intervals. For Lebesgue measure to be truly useful, it must be extended from $\mathcal{R}$ to this σ-algebra, whose members are called the **Borel sets** of $\mathbb{R}^1$. Borel sets of $\mathbb{R}^1$ can be fairly complicated. Each open set is a Borel set because it is the countable union of bounded open intervals. Each closed set is a Borel set, being the complement of an open set, and each compact set is a Borel set because compact subsets of $\mathbb{R}^1$ are closed. In addition, any countable set, such as the set $\mathbb{Q}$ of rationals, is a Borel set as the countable union of one-point sets.

The extension is carried out by the general Extension Theorem that will be stated now and will be proved in Section 5. The theorem gives both existence and uniqueness for an extension, but not without an additional hypothesis. The need for an additional hypothesis to ensure uniqueness is closely related to the need to assume some finiteness condition on ρ in Corollary 5.3: even though each member of a decreasing sequence of sets has infinite measure, the intersection of the sets need not have infinite measure. To see what can go wrong for the Extension Theorem, consider the ring $\mathcal{R}'$ of subsets of $\mathbb{R}^1$ consisting of all finite unions of bounded intervals with rational endpoints; the individual intervals may or may not contain their endpoints. If a set function μ is defined on this ring by assigning to each set the number of elements in the set, then μ is completely additive. Each interval in $\mathbb{R}^1$ can be obtained as the union of two sets—a countable union of intervals with rational endpoints and a countable intersection of intervals with rational endpoints. It follows that the smallest σ-ring containing $\mathcal{R}'$ is the σ-algebra of all Borel sets. The set function μ can be extended to the Borel sets in more than one way. In fact, each one-point set consisting of a rational must get measure 1, but a one-point set consisting of an irrational can be assigned any measure.

The additional hypothesis for the Extension Theorem is that the given nonnegative completely additive set function ν on a ring of sets $\mathcal{R}$ be σ**-finite**, i.e., that any member of $\mathcal{R}$ be contained in the countable union of members of $\mathcal{R}$ on which ν is finite. An obvious sufficient condition for σ-finiteness is that $\nu(E)$ be finite for every set in $\mathcal{R}$. This sufficient condition is satisfied by Lebesgue measure on the elementary sets, and thus the theorem proves that Lebesgue measure extends in a unique fashion to be a measure on the Borel sets.

The condition of σ-finiteness is less restrictive than a requirement that X be the countable union of sets in $\mathcal{R}$ of finite measure, another condition that is satisfied in the case of Lebesgue measure. The condition of σ-finiteness on a ring allows for some very large measures when all the sets are in a sense generated by the sets of finite measure. For example, if $\mathcal{R}$ is the ring of finite subsets of an uncountable set and ν is the counting measure, the σ-finiteness condition is satisfied. In most areas of mathematics, these very large measures rarely arise.

Theorem 5.5 (Extension Theorem). Let $\mathcal{R}$ be a ring of subsets of a nonempty set X, and let ν be a nonnegative completely additive set function on $\mathcal{R}$ that is σ-finite on $\mathcal{R}$. Then ν extends uniquely to a measure μ on the smallest σ-ring containing $\mathcal{R}$.

A **measure space** is defined to be a triple $(X, \mathcal{A}, \mu)$, where X is a nonempty set, $\mathcal{A}$ is a σ-algebra of subsets of X, and μ is a measure on X. The measure space is **finite** if $\mu(X) < +\infty$; it is **σ-finite** if X is the countable union of sets on which μ is finite. The real line, together with the σ-algebra of Borel sets and Lebesgue measure, is a σ-finite measure space.

2. Measurable Functions

In this section, X denotes a nonempty set, and $\mathcal{A}$ is a σ-algebra of subsets of X. The **measurable sets** are the members of $\mathcal{A}$.

We say that a function $f : X \to \mathbb{R}^*$ is **measurable** if

(i) $f^{-1}([-\infty, c))$ is a measurable set for every real number c.

Equivalently the measurability of f may be defined by any of the following conditions:

(ii) $f^{-1}([-\infty, c])$ is a measurable set for every real number c,
(iii) $f^{-1}((c, +\infty])$ is a measurable set for every real number c,
(iv) $f^{-1}([c, +\infty])$ is a measurable set for every real number c.

In fact, the implications (i) implies (ii), (ii) implies (iii), (iii) implies (iv), and (iv) implies (i) follow from the identities[4]

$$\begin{aligned}
f^{-1}([-\infty, c]) &= \bigcap_{n=1}^{\infty} f^{-1}([-\infty, c + \tfrac{1}{n})),\\
f^{-1}((c, +\infty]) &= (f^{-1}([-\infty, c]))^c,\\
f^{-1}([c, +\infty]) &= \bigcap_{n=1}^{\infty} f^{-1}((c - \tfrac{1}{n}, +\infty]),\\
f^{-1}([-\infty, c)) &= (f^{-1}([c, +\infty]))^c.
\end{aligned}$$

EXAMPLES.

(1) If $\mathcal{A} = \{\varnothing, X\}$, then only the constant functions are measurable.

(2) If $\mathcal{A}$ consists of all subsets of X, then every function from X to $\mathbb{R}^*$ is measurable.

[4]Manipulations with inverse images of sets are discussed in Section A1 of Appendix A.

(3) If $X = \mathbb{R}^1$ and $\mathcal{A}$ consists of the Borel sets of $\mathbb{R}^1$, the measurable functions are often called **Borel measurable**. Every continuous function is Borel measurable by (i) because the inverse image of every open set is open. Any function that is 1 on an open or compact set and is 0 off that set is Borel measurable. It is shown in Problem 33 at the end of the chapter that not every Riemann integrable function (when set equal to 0 off some bounded interval) is Borel measurable. However, let us verify that every function that is continuous except at countably many points is Borel measurable. In fact, let C be the exceptional countable set. The restriction of f to the metric space $\mathbb{R} - C$ is continuous, and hence the inverse image in $\mathbb{R} - C$ of any open set $[-\infty, c)$ is open in $\mathbb{R} - C$. Hence the inverse image is the countable union of sets $(a, b) - C$, and these are Borel sets. The full inverse image in $\mathbb{R}$ of $[-\infty, c)$ under f is the union of a countable set and this subset of $\mathbb{R} - C$ and hence is a Borel set.

(4) If $X = \mathbb{R}^1$ and if $\mathcal{A}$ consists of the "Lebesgue measurable sets" in a sense to be defined in Section 5, the measurable functions are often called **Lebesgue measurable**. Every Borel measurable function is Lebesgue measurable, and so is every Riemann integrable function (when set equal to 0 off some bounded interval).

The next proposition discusses, among other things, functions f^+, f^-, and $|f|$ defined by $f^+(x) = \max\{f(x), 0\}$, $f^-(x) = -\min\{f(x), 0\}$, and $|f|(x) = |f(x)|$. Then $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

Proposition 5.6.

(a) Constant functions are always measurable.

(b) If f is measurable, then the inverse image of any interval is measurable.

(c) If f is measurable, then the inverse image of any open set in $\mathbb{R}^*$ is measurable.

(d) If f is measurable, then the functions f^+, f^-, and $|f|$ are measurable.

PROOF. In (a), the inverse image of a set under a constant function is either $\varnothing$ or X and in either case is measurable. In (b), the inverse image of an interval is the intersection of two sets of the kind described in (i) through (iv) above and hence is measurable. In (c), any open set in $\mathbb{R}^*$ is the countable union of open intervals, and the measurability of the inverse image follows from (b) and the closure of $\mathcal{A}$ under countable unions. In (d), $(f^+)^{-1}((c, +\infty))$ equals $f^{-1}((c, +\infty))$ if $c \geq 0$ and equals X if $c < 0$. The measurability of f^- and $|f|$ are handled similarly. □

Next we deal with measurability of sums and products, allowing for values $+\infty$ and $-\infty$. Recall from Section I.1 that multiplication is everywhere defined in $\mathbb{R}^*$ and that the product in $\mathbb{R}^*$ of 0 with anything is 0.

Proposition 5.7. Let f and g be measurable functions, and let a be in $\mathbb{R}$. Then af and fg are measurable, and $f+g$ is measurable provided the sum $f(x)+g(x)$ is everywhere defined.

PROOF. For $f+g$, with $\mathbb{Q}$ denoting the rationals,

$$(f+g)^{-1}(c,+\infty] = \bigcup_{r\in\mathbb{Q}} f^{-1}(c+r,+\infty] \cap g^{-1}(-r,+\infty].$$

If $a=0$, then $af=0$, and 0 is measurable. If $a\neq 0$, then

$$(af)^{-1}(c,+\infty] = \begin{cases} f^{-1}\big(\frac{c}{a},+\infty\big] & \text{if } a>0, \\ f^{-1}\big[-\infty,\frac{c}{a}\big) & \text{if } a<0. \end{cases}$$

If f and g are measurable and are ≥ 0, then

$$(fg)^{-1}(c,+\infty] = \begin{cases} \bigcup_{r\in\mathbb{Q},\,r>0} f^{-1}\big(\frac{c}{r},+\infty\big] \cap g^{-1}(r,+\infty] & \text{if } c\geq 0, \\ X & \text{if } c<0. \end{cases}$$

Hence fg is measurable in this special case. In the general case the formula $fg = f^+g^+ + f^-g^- - f^+g^- - f^-g^+$ exhibits fg as the everywhere-defined sum of measurable functions. $\square$

Proposition 5.8. If $\{f_n\}$ is a sequence of measurable functions, then the functions

(a) $\sup_n f_n$,
(b) $\inf_n f_n$,
(c) $\limsup_n f_n$,
(d) $\liminf f_n$,

are all measurable.

PROOF. For (a) and (b), we have $(\sup f_n)^{-1}(c,+\infty] = \bigcup_{n=1}^\infty f_n^{-1}(c,+\infty]$ and $(\inf f_n)^{-1}([-\infty,c) = \bigcup_{n=1}^\infty f_n^{-1}[-\infty,c)$. For (c) and (d), we have $\limsup_n f_n = \inf_n \sup_{k\geq n} f_k$ and $\liminf_n f_n = \sup_n \inf_{k\geq n} f_k$. $\square$

Corollary 5.9. The pointwise maximum and the pointwise minimum of a finite set of measurable functions are both measurable.

PROOF. These are special cases of (a) and (b) in the proposition. $\square$

Corollary 5.10. If $\{f_n\}$ is a sequence of measurable functions and if $f(x) = \lim f_n(x)$ exists in $\mathbb{R}^*$ at every x, then f is measurable.

PROOF. This is the special case of (c) and (d) in the proposition in which $\limsup_n f_n = \liminf_n f_n$. $\square$

The above results show that the set of measurable functions is closed under pointwise limits, as well as the arithmetic operations and max and min. Since the measurable functions will be the ones we attempt to integrate, we can hope for good limit theorems from Lebesgue integration, as well as the familiar results about arithmetic operations and ordering properties.

If E is a subset of X, the **indicator function**[5] I_E of E is the function that is 1 on E and is 0 elsewhere. The set $(I_E)^{-1}(c, +\infty]$ is $\varnothing$ or E or X, depending on the value of c. Therefore I_E is a measurable function if and only if E is a measurable set.

A **simple function** $s : X \to \mathbb{R}^*$ is a function s with finite image contained in $\mathbb{R}$. Every simple function s has a unique representation as $s = \sum_{n=1}^{N} c_n I_{E_n}$, where the c_n are distinct real numbers and the E_n are disjoint nonempty sets with union X. In fact, the set of numbers c_n equals the image of s, and E_n is the set where s takes the value c_n. This expansion of s will be called the **canonical expansion** of s. The set $s^{-1}(c, +\infty]$ is the union of the sets E_n such that $c < c_n$, and it follows that s is a measurable function if and only if all of the sets E_n in the canonical expansion are measurable sets.

Proposition 5.11. For any function $f : X \to [0, +\infty]$, there exists a sequence of simple functions $s_n \geq 0$ with the property that for each x in X, $\{s_n(x)\}$ is a monotone increasing sequence in $\mathbb{R}$ with limit $f(x)$ in $\mathbb{R}^*$. If f is measurable, then the simple functions s may be taken to be measurable.

PROOF. For $1 \leq n < \infty$ and $1 \leq j \leq n2^n$, let

$$E_{nj} = f^{-1}\Big[\frac{j-1}{2^n}, \frac{j}{2^n}\Big), \quad F_n = f^{-1}[n, +\infty), \quad s_n = \sum_{j=1}^{n2^n} \frac{j-1}{2^n} I_{E_{nj}} + n I_{F_n}.$$

Then $\{s_n\}$ has the required properties. $\square$

By convention from now on, *simple functions will always be understood to be measurable*.

3. Lebesgue Integral

Throughout this section, $(X, \mathcal{A}, \mu)$ denotes a measure space. The measurable sets continue to be those in $\mathcal{A}$. Our objective in this section is to define the Lebesgue

[5] As noted in Chapter III, indicator functions are called "characteristic functions" by many authors, but the term "characteristic function" has another meaning in probability theory and is best avoided as a substitute for "indicator function" in any context where probability might play a role.

integral. We defer any systematic discussion of properties of the integral to Section 4.

Just as with the Riemann integral, the Lebesgue integral is defined by means of an approximation process. In the case of the Riemann integral, the process is to use upper sums and lower sums, which capture an approximate value of an integral by adding contributions influenced by proximity in the *domain* of the integrand. The process is qualitatively different for the Lebesgue integral, which captures an approximate value of an integral by adding contributions based on what happens in the *image* of the integrand.

Let s be a simple function ≥ 0. By our convention at the end of the previous section, we have incorporated measurability into the definition of simple function. Let E be a measurable set, and let $s = \sum_{n=1}^{N} c_n I_{A_n}$ be the canonical expansion of s. We define $\mathcal{I}_E(s) = \sum_{n=1}^{N} c_n \mu(A_n \cap E)$. This kind of object will be what we use as an approximation in the definition of the Lebesgue integral; the formula shows the sense in which $\mathcal{I}_E(s)$ is built from the *image* of the integrand.

If $f \geq 0$ is a measurable function and E is a measurable set, we define the **Lebesgue integral** of f on the set E with respect to the measure μ to be

$$\int_E f\,d\mu = \int_E f(x)\,d\mu(x) = \sup_{\substack{0 \leq s \leq f, \\ s \text{ simple}}} \mathcal{I}_E(s).$$

This is well-defined as a member of $\mathbb{R}^*$ without restriction as long as E is a measurable set and the measurable function f is ≥ 0 everywhere on X. It is evident in this case that $\int_E f\,d\mu \geq 0$ and that $\int_E 0\,d\mu = 0$.

For a general measurable function f, not necessarily ≥ 0, the integral may or may not be defined. We write $f = f^+ - f^-$. The functions f^+ and f^- are ≥ 0 and are measurable by Proposition 5.6d, and consequently $\int_E f^+\,d\mu$ and $\int_E f^-\,d\mu$ are well-defined members of $\mathbb{R}^*$. If $\int_E f^+\,d\mu$ and $\int_E f^-\,d\mu$ are not both infinite, then we define

$$\int_E f\,d\mu = \int_E f(x)\,d\mu(x) = \int_E f^+\,d\mu - \int_E f^-\,d\mu.$$

This definition is consistent with the definition in the special case $f \geq 0$, since such an f has $f^- = 0$ and therefore $\int_E f^-\,d\mu = 0$. We say that f is **integrable** if $\int_E f^+\,d\mu$ and $\int_E f^-\,d\mu$ are *both* finite. In this case the subsets of E where f is $+\infty$ and where f is $-\infty$ have measure 0. In fact, if S is the subset of E where f^+ is $+\infty$, then the inequality $\int_E f^+\,d\mu \geq \mathcal{I}_E(CI_S) = C\mu(S)$ for every $C > 0$ shows that $\mu(S) \leq C^{-1}\int_E f^+\,d\mu$ for every C; hence $\mu(S) = 0$. A similar argument applies to the set where f^- is $+\infty$.

We shall give some examples of integration after showing that the definition of $\int_E f\,d\mu$ reduces to $\mathcal{I}_E(f)$ if f is nonnegative and simple. The first lemma below will make use of the additivity of μ, and the second lemma will make use of the fact that μ is nonnegative.

Lemma 5.12. Let $s = \sum_{n=1}^{N} c_n I_{A_n}$ be the canonical expansion of a simple function ≥ 0, and let $s = \sum_{m=1}^{M} d_n I_{B_m}$ be another expansion in which the d_m are ≥ 0 and the B_m are disjoint and measurable. Then $\mathcal{I}_E(s) = \sum_{m=1}^{M} d_m \mu(B_m \cap E)$.

PROOF. Adjoin the term $0 \cdot I_{(\bigcup_m B_m)^c}$ to the second expansion, if necessary, to make $\bigcup_{m=1}^{M} B_m = X$. Without loss of generality, we may assume that no B_m is empty. Then the fact that the sets B_m are disjoint and nonempty with union X implies that the image of s is $\{d_1, \dots, d_M\}$. Thus we can write $d_m = c_{n(m)}$ for each m. Since $A_n = s^{-1}(\{c_n\})$, we see that $B_m \subseteq A_{n(m)}$. Since the B_m are disjoint with union X, we obtain

$$A_k = \bigcup_{\{m \mid n(m)=k\}} B_m$$

disjointly. The additivity of μ gives $\mu(A_k \cap E) = \sum_{\{m \mid n(m)=k\}} \mu(B_m \cap E)$, and thus $c_k \mu(A_k \cap E) = \sum_{\{m \mid n(m)=k\}} d_m \mu(B_m \cap E)$. Summing on k, we obtain the conclusion of the lemma. $\square$

Lemma 5.13. If s and t are nonnegative simple functions and if $t \leq s$ on E, then $\mathcal{I}_E(t) \leq \mathcal{I}_E(s)$.

PROOF. If $s = \sum_{j=1}^{J} c_j I_{A_j}$ and $t = \sum_{k=1}^{K} d_k I_{B_k}$ are the canonical expansions of s and t, then $\bigcup_{j,k} (A_j \cap B_k) = X$ disjointly. Hence we can write

$$s = \sum_{j,k} c_j I_{A_j \cap B_k} \qquad \text{and} \qquad t = \sum_{j,k} d_k I_{A_j \cap B_k}.$$

Lemma 5.12 shows that

$$\mathcal{I}_E(s) = \sum_{j,k} c_j \mu(A_j \cap B_k \cap E) \qquad \text{and} \qquad \mathcal{I}_E(t) = \sum_{j,k} d_k \mu(A_j \cap B_k \cap E).$$

We now have term-by-term inequality: either $\mu(A_j \cap B_k \cap E) = 0$ for a term, or $A_j \cap B_k \cap E \neq \varnothing$ and any x in $A_j \cap B_k \cap E$ has $t(x) \leq s(x)$ and exhibits $d_k \leq c_j$. $\square$

Proposition 5.14. If $s \geq 0$ is a simple function, then $\int_E s\,d\mu = \mathcal{I}_E(s)$ for every measurable set E.

PROOF. If t is a simple function with $0 \leq t \leq s$ everywhere, then Lemma 5.13 gives $\mathcal{I}_E(t) \leq \mathcal{I}_E(s)$. Hence $\int_E s\,d\mu = \sup_{0 \leq t \leq s} \mathcal{I}_E(t) \leq \mathcal{I}_E(s)$. On the other hand, we certainly have $\mathcal{I}_E(s) \leq \sup_{0 \leq t \leq s} \mathcal{I}_E(t) = \int_E s\,d\mu$, and thus $\int_E s\,d\mu = \mathcal{I}_E(s)$. $\square$

EXAMPLES.

(1) Let $\mathcal{A} = \{\varnothing, X\}$ and $\mu(X) = 1$. Only the constant functions are measurable, and $\int_\varnothing c\, d\mu = 0$ and $\int_X c\, d\mu = c$.

(2) Let X be a nonempty countable set, let $\mathcal{A}$ consist of all subsets of X, and let μ be defined by nonnegative finite weights w_i attached to each point i in X. If $f = \{f_i\}$ is a real-valued function, then the integral of f over X is $\sum f_i w_i$ provided the integrals of f^+ and f^- are not both infinite, i.e., provided every rearrangement of the series $\sum f_i w_i$ converges in $\mathbb{R}^*$ to the same sum. By contrast, f is integrable if and only if the series $\sum f_i w_i$ is absolutely convergent; this is a stronger condition since the sum has to be in $\mathbb{R}$. In the special case that all the weights w_i are 1, the theory of the Lebesgue integral over X reduces to the theory of infinite series for which every rearrangement of the series converges in $\mathbb{R}^*$ to the same sum. This is a very important special case for testing the validity of general assertions about Lebesgue integration.

(3) Let $(X, \mathcal{A}, \mu)$ be the real line $\mathbb{R}^1$ with $\mathcal{A}$ consisting of the Borel sets and with μ equal to Lebesgue measure m. Recall that real-valued continuous functions on $\mathbb{R}^1$ are measurable. For such a function f, the assertion is that

$$\int_{[a,x)} f\, dm = \int_a^x f(t)\, dt,$$

the left side being a Lebesgue integral and the right side being a Riemann integral. Proving this assertion involves using some properties of the Lebesgue integral that will be proved in the next section. We give the argument now before these properties have been established, in order to emphasize the importance of each of these properties: If $h > 0$, then

$$\begin{aligned}\frac{1}{h}\Big[\int_{[a,x+h)} f\, dm - \int_{[a,x)} f\, dm\Big] - f(x) &= \frac{1}{h}\int_{[x,x+h)} f\, dm - f(x)\\ &= \frac{1}{h}\int_{[x,x+h)} [f - f(x)]\, dm.\end{aligned}$$

The absolute value of the left side is then

$$\begin{aligned}\leq \frac{1}{h}\int_{[x,x+h)} |f - f(x)|\, dm &\leq \frac{1}{h}\sup_{t\in[x,x+h)} |f(t) - f(x)|\, m([x, x+h))\\ &= \sup_{t\in[x,x+h)} |f(t) - f(x)|,\end{aligned}$$

and the right side tends to 0 as h decreases to 0, by continuity of f at x. If $h < 0$, then the argument corresponding to the first display is

$$\begin{aligned}\frac{1}{h}\Big[\int_{[a,x+h)} f\, dm - \int_{[a,x)} f\, dm\Big] - f(x) &= -\frac{1}{h}\int_{[x-|h|,x)} f\, dm - f(x)\\ &= \frac{1}{|h|}\int_{[x-|h|,x)} [f - f(x)]\, dm.\end{aligned}$$

The absolute value of the left side is then $\le \sup_{t\in[x-|h|,x)} |f(t) - f(x)|$, and this tends to 0 as h increases to 0, by continuity of f at x. We conclude that $\int_{[a,\cdot)} f\,dm$ is differentiable with derivative f. By the Fundamental Theorem of Calculus for the Riemann integral, together with a corollary of the Mean Value Theorem, $\int_{[a,x)} f\,dm = \int_a^x f(t)\,dt + c$ for all x and some constant c. Putting $x = a$, we see that $c = 0$. Therefore the Riemann and Lebesgue integrals coincide for continuous functions on bounded intervals $[a, b)$.

4. Properties of the Integral

In this section, $(X, \mathcal{A}, \mu)$ continues to denote a measure space. Our objective is to establish basic properties of the Lebesgue integral, including properties that indicate how Lebesgue integration interacts with passages to the limit. The properties that we establish will include all remaining properties needed to justify the argument in Example 3 at the end of the previous section.

Proposition 5.15. The Lebesgue integral has these four properties:

(a) If f is a measurable function and $\mu(E) = 0$, then $\int_E f\,d\mu = 0$.

(b) If E and F are measurable sets with $F \subseteq E$ and if f is a measurable function, then $\int_F f^+\,d\mu \le \int_E f^+\,d\mu$ and $\int_F f^-\,d\mu \le \int_E f^-\,d\mu$. Consequently, if $\int_E f\,d\mu$ is defined, then so is $\int_F\,d\mu$.

(c) If c is a constant function with its value in $\mathbb{R}^*$, then $\int_E c\,d\mu = c\mu(E)$.

(d) If $\int_E f\,d\mu$ is defined and if c is in $\mathbb{R}$, then $\int_E cf\,d\mu$ is defined and $\int_E cf\,d\mu = c\int_E f\,d\mu$. If f is integrable on E, then so is cf.

PROOF. In (a), it is enough to deal with f^+ and f^- separately, and then it is enough to handle $s \ge 0$ simple. For such an s, Proposition 5.14 says that the integral equals $\mathcal{I}_E(s)$, and the definition shows that this is 0. In (b), Proposition 5.14 makes it clear that the inequalities are valid for any simple function ≥ 0, and then the general case follows by taking the supremum first for $0 \le s \le f^+$ and then for $0 \le s \le f^-$. In (c), if $0 \le c < +\infty$, then c is simple, and the integral equals $\mathcal{I}_E(c) = c\mu(E)$ by Proposition 5.14. If $c = +\infty$, then the case $\mu(E) = 0$ follows from (a) and the case $\mu(E) > 0$ is handled by the observations that $\int_E c\,d\mu \ge \mathcal{I}_E(n) = n\mu(E)$ and that the right side tends to $+\infty$ as n tends to $+\infty$. For $c \le 0$, we have $\int_E c\,d\mu = -\int_E(-c)\,d\mu$ by definition, and then the result follows from the previous cases. In (d), we may assume, without loss of generality, that $f \ge 0$ and $c \ge 0$. Then $\int_E cf\,d\mu = \sup_{0\le s\le cf} \mathcal{I}_E(s) = \sup_{0\le ct\le cf} \mathcal{I}_E(ct) = c\sup_{0\le t\le f} \mathcal{I}_E(t) = c\int_E f\,d\mu$, and (d) is proved. □

Proposition 5.16. If f and g are measurable functions, if their integrals over E are defined, and if $f(x) \le g(x)$ on E, then $\int_E f\,d\mu \le \int_E g\,d\mu$.

REMARK. Observe that the inequality $f(x) \le g(x)$ is assumed only on E, despite the definitions that take into account values of a function everywhere on X. This "localization" property of the integral is as one wants it to be.

PROOF. First suppose that $f \ge 0$ and $g \ge 0$. If s is any simple function with $0 \le s \le f$, define t to equal s on E and to equal 0 off E. Then $0 \le t \le g$, and Lemma 5.13 gives $\mathcal{I}_E(s) = \mathcal{I}_E(t) \le \int_E g\,d\mu$. Hence $\int_E f\,d\mu \le \int_E g\,d\mu$ when $f \ge 0$ and $g \ge 0$.

In the general case the inequality $f(x) \le g(x)$ on E implies that $f^+(x) \le g^+(x)$ on E and $f^-(x) \ge g^-(E)$ on E. The special case gives $\int_E f^+\,d\mu \le \int_E g^+\,d\mu$ and $\int_E f^-\,d\mu \ge \int_E g^-\,d\mu$. Subtracting these inequalities, we obtain the desired result. □

Corollary 5.17. If f and g are measurable functions that are equal on E and if $\int_E f\,d\mu$ is defined, then $\int_E g\,d\mu$ is defined and $\int_E f\,d\mu = \int_E g\,d\mu$.

PROOF. Apply Proposition 5.16 to the following inequalities on E, and then sort out the results: $f^+ \le g^+$, $f^+ \ge g^+$, $f^- \le g^-$, and $f^- \ge g^-$. □

Corollary 5.18. If f is a measurable function, then f is integrable on E if either

(a) there is a function g integrable on E such that $|f(x)| \le g(x)$ on E, or
(b) $\mu(E)$ is finite and there is a real number c such that $|f(x)| \le c$ on E.

PROOF. For (a), apply Proposition 5.16 to the inequalities $f^+ \le g$ and $f^- \le g$ valid on E. For (b), use the formula for $\int_E c\,d\mu$ in Proposition 5.15c and apply (a). □

We turn our attention now to properties that indicate how Lebesgue integration interacts with passages to the limit. These make essential use of the complete additivity of the measure μ. We shall bring this hypothesis to bear initially through the following theorem.

Theorem 5.19. Let f be a fixed measurable function, and suppose that $\int_X f\,d\mu$ is defined. Then the set function $\rho(E) = \int_E f\,d\mu$ is completely additive.

PROOF. We have $\rho(\varnothing) = 0$ by Proposition 5.15a, since $\mu(\varnothing) = 0$. We shall prove that if $f \ge 0$, then ρ is completely additive. The general case follows from this by applying the result to f^+ and f^- separately and by using the fact that $\int_X f^+\,d\mu$ and $\int_X f^-\,d\mu$ are not both infinite. Thus we are to show that if $E = \bigcup_{n=1}^\infty E_n$ disjointly and if $f \ge 0$, then $\rho(E) = \sum_{n=1}^\infty \rho(E_n)$.

For simple $s \ge 0$ with canonical expansion $s = \sum_{n=1}^N c_n I_{A_n}$, the identity $\mathcal{I}_F(s) = \sum_{n=1}^N c_n \mu(A_n \cap F)$ and the complete additivity of μ show that $\mathcal{I}_F(s)$ is a completely additive function of the set F. Thus for s simple with $0 \le s \le f$, we have

$$\mathcal{I}_E(s) = \sum_{n=1}^{\infty} \mathcal{I}_{E_n}(s) \le \sum_{n=1}^{\infty} \rho(E_n).$$

Hence

$$\rho(E) = \sup_{0 \le s \le f} \mathcal{I}_E(s) \le \sum_{n=1}^{\infty} \rho(E_n).$$

We now prove the reverse inequality. By Proposition 5.15b, $\rho(E) \ge \rho(E_n)$ for every n, since $f = f^+$. Hence if $\rho(E_n) = +\infty$ for any n, the desired result is proved. Thus assume that $\rho(E_n) < +\infty$ for all n. Let $\epsilon > 0$ be given, and choose simple functions t and u that are ≥ 0 and are $\le f$ and have

$$\mathcal{I}_{E_1}(t) \ge \int_{E_1} f \, d\mu - \epsilon \qquad \text{and} \qquad \mathcal{I}_{E_2}(u) \ge \int_{E_2} f \, d\mu - \epsilon.$$

Let s be the pointwise maximum $s = \max\{t, u\}$. Then s is simple, and Lemma 5.13 gives $\mathcal{I}_{E_1}(s) \ge \mathcal{I}_{E_1}(t)$ and $\mathcal{I}_{E_2}(s) \ge \mathcal{I}_{E_2}(u)$. Consequently

$$\begin{aligned} \rho(E_1 \cup E_2) &= \int_{E_1 \cup E_2} f \, d\mu \ge \mathcal{I}_{E_1 \cup E_2}(s) = \mathcal{I}_{E_1}(s) + \mathcal{I}_{E_2}(s) \\ &\ge \mathcal{I}_{E_1}(t) + \mathcal{I}_{E_2}(u) \ge \int_{E_1} f \, d\mu + \int_{E_2} f \, d\mu - 2\epsilon \\ &= \rho(E_1) + \rho(E_2) - 2\epsilon. \end{aligned}$$

Since ϵ is arbitrary, $\rho(E_1 \cup E_2) \ge \rho(E_1) + \rho(E_2)$. By induction, we obtain $\rho(E_1 \cup \cdots \cup E_n) \ge \rho(E_1) + \cdots + \rho(E_n)$ for every n, and thus $\rho(E) \ge \rho(E_1) + \cdots + \rho(E_n)$ by another application of Proposition 5.15b. Therefore $\rho(E) \ge \sum_{n=1}^{\infty} \rho(E_n)$, and the reverse inequality has been proved. □

We give five corollaries that are consequences of Corollary 5.17 and Theorem 5.19. The first three make use only of additivity, not of complete additivity.

Corollary 5.20. If $\int_E f \, d\mu$ is defined, then $\int_X I_E f \, d\mu$ is defined and equals $\int_E f \, d\mu$.

PROOF. It is sufficient to handle f^+ and f^- separately. Then both integrals are defined, and $\int_E f \, d\mu = \int_E I_E f \, d\mu + \int_{E^c} 0 \, d\mu = \int_E I_E f \, d\mu + \int_{E^c} I_E f \, d\mu = \int_X I_E f \, d\mu$. □

Corollary 5.21. If $\int_E f\,d\mu$ is defined, then $\big|\int_E f\,d\mu\big| \le \int_E |f|\,d\mu$. If f is integrable on E, so is $|f|$.

PROOF. Let $E_1 = E \cap f^{-1}([0,+\infty])$ and $E_2 = E \cap f^{-1}([-\infty,0))$. Then use of the triangle inequality gives

$$\begin{aligned}\big|\textstyle\int_E f\,d\mu\big| &= \big|\textstyle\int_{E_1} f^+\,d\mu - \int_{E_2} f^-\,d\mu\big| \le \textstyle\int_{E_1} f^+\,d\mu + \int_{E_2} f^-\,d\mu \\ &= \textstyle\int_{E_1} |f|\,d\mu + \int_{E_2} |f|\,d\mu = \int_E |f|\,d\mu.\end{aligned}$$

If f is integrable on E, both $\int_{E_1} f^+\,d\mu$ and $\int_{E_2} f^-\,d\mu$ are finite. Their sum is $\int_E |f|\,d\mu$. $\square$

Corollary 5.22. If f is a measurable function and $\mu(E\,\Delta\,F) = 0$, then $\int_E f\,d\mu = \int_F f\,d\mu$, provided one of the integrals exists.

PROOF. Without loss of generality, we may assume that $f \ge 0$. Then both integrals are defined. Since $E\,\Delta\,F = (E-F)\cup(F-E)$, we have $\mu(E-F) = \mu(F-E) = 0$. Then Theorem 5.19 and Proposition 5.15a give $\int_E f\,d\mu = \int_{E-F} f\,d\mu + \int_{E\cap F} f\,d\mu = 0 + \int_{E\cap F} f\,d\mu = \int_{F-E} f\,d\mu + \int_{E\cap F} f\,d\mu = \int_F f\,d\mu$. $\square$

Corollary 5.23. If f is a measurable function and if the set $A = \{x \mid f(x) \ne 0\}$ has $\mu(A) = 0$, then $\int_X f\,d\mu = 0$. Conversely if f is measurable, is ≥ 0, and has $\int_X f\,d\mu = 0$, then $A = \{x \mid f(x) \ne 0\}$ has $\mu(A) = 0$.

REMARKS. When a set where some condition fails to hold has measure 0, one sometimes says that the condition holds **almost everywhere**, or **a.e.**, or at **almost every** point. If there is any ambiguity about what measure is being referred to, one says "a.e. [$d\mu$]." Thus the conclusion in the converse half of the above proposition is that f is zero a.e. [$d\mu$].

PROOF. For the first statement, Corollary 5.20 gives $\int_X f\,d\mu = \int_X I_A f\,d\mu = \int_A f\,d\mu = 0$. Conversely let $A_n = f^{-1}\big(\big[\tfrac{1}{n},+\infty\big]\big)$. This is a measurable set. Since f is ≥ 0, $A = \bigcup_{n=1}^\infty A_n$. Proposition 5.1g and complete additivity of μ give $\mu(A) \le \sum_{n=1}^\infty \mu(A_n)$. If $\mu(A_n) > 0$ for some n, then $\int_X f\,d\mu = \int_{A_n} f\,d\mu + \int_{A_n^c} f\,d\mu \ge \int_{A_n} \frac{1}{n}\,d\mu = \frac{1}{n}\,\mu(A_n) > 0$, and we obtain a contradiction. We conclude that $\mu(A_n) = 0$ for all n and hence that $\mu(A) = 0$. $\square$

Corollary 5.24. If $f \ge 0$ is an integrable function on X, then for any $\epsilon > 0$, there exists a $\delta > 0$ such that $\int_E f\,d\mu \le \epsilon$ for every measurable set E with $\mu(E) \le \delta$.

PROOF. Let $\epsilon > 0$ be given. If $N > 0$ is an integer, then the sets $S_N = \{x \in X \mid f(x) \geq N\}$ form a decreasing sequence whose intersection is $S = \{x \in X \mid f(x) = +\infty\}$. Since f is integrable, $\mu(S) = 0$ and therefore $\int_S f\,d\mu = 0$. The finiteness of $\int_X f\,d\mu$, together with Corollary 5.3 and the complete additivity of $E \mapsto \int_E f\,d\mu$ given in Theorem 5.19, implies that $\lim_N \int_{S_N} f\,d\mu = 0$. Choose N large enough so that $\int_{S_N} f\,d\mu \leq \epsilon/2$, and then choose $\delta = \epsilon/(2N)$. If $\mu(E) \leq \delta$, then

$$\begin{aligned}\int_E f\,d\mu &= \int_{S_N \cap E} f\,d\mu + \int_{S_N^c \cap E} f\,d\mu \\ &\leq \int_{S_N} f\,d\mu + \int_{S_N^c \cap E} N\,d\mu \leq \epsilon/2 + N\mu(E) \leq \epsilon/2 + \epsilon/2 = \epsilon,\end{aligned}$$

and the proof is complete. □

In a number of the remaining results in the section, a sequence $\{f_n\}$ of measurable functions converges pointwise to a function f. Corollary 5.10 assures us that f is measurable. Suppose that $\int_E f_n\,d\mu$ exists for each n. Is it true that $\int_E f\,d\mu$ exists, is it true that $\lim_n \int_E f_n\,d\mu$ exists, and if both exist, are they equal? Once again we encounter an interchange-of-limits problem, and there is no surprise from the general fact: all three answers can be "no" in particular cases. Examples of the failure of the limit of the integral to equal the integral of the limit are given below. After giving the examples, we shall discuss theorems that give "yes" answers under additional hypotheses.

EXAMPLES.

(1) Let X be the set of positive integers, let $\mathcal{A}$ consist of all subsets of X, and let μ be counting measure. A measurable function f is a sequence $\{f(k)\}$ with values in $\mathbb{R}^*$. Define a sequence $\{f_n\}$ of measurable functions for $n \geq 1$ by taking

$$f_n(k) = \begin{cases} 1/n & \text{if } k \leq n, \\ 0 & \text{if } k > n. \end{cases}$$

Then $\int_X f_n\,d\mu = 1$ for all n, $\lim f_n = 0$ pointwise, and

$$\int_X \lim f_n\,d\mu < \lim \int_X f_n\,d\mu.$$

(2) Let the measure space be $X = \mathbb{R}^1$ with the Borel sets and Lebesgue measure m. Define

$$f_n(x) = \begin{cases} n & \text{for } 0 < x < 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

Then the same phenomenon results, and everything of interest is taking place within [0, 1]. So the difficulty in the previous example does not result from the fact that X has infinite measure.

Theorem 5.25 (Monotone Convergence Theorem). Let E be a measurable set, and suppose that $\{f_n\}$ is a sequence of measurable functions that satisfy

$$0 \le f_1(x) \le f_2(x) \le \cdots \le f_n(x) \le \cdots$$

for all x. Put $f(x) = \lim_n f_n(x)$, the limit being taken in $\mathbb{R}^*$. Then $\int_E f\,d\mu$ and $\lim_n \int_E f_n\,d\mu$ both exist, and

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

REMARKS. This theorem generalizes Corollary 1.14, which is the special case of the Monotone Convergence Theorem in which X is the set of positive integers, every subset is measurable, and μ is counting measure. In the general setting of the Monotone Convergence Theorem, one of the by-products of the theorem is that we obtain an easier way of dealing with the definition of $\int_E f\,d\mu$ for $f \ge 0$. Instead of using the totality of simple functions between 0 and f, we may use a single increasing sequence with pointwise limit f, such as the one given by Proposition 5.11. The proof of Proposition 5.26 below will illustrate how we can take advantage of this fact.

PROOF. Since f is the pointwise limit of measurable functions and is ≥ 0, f is measurable and $\int_E f\,d\mu$ exists in $\mathbb{R}^*$. Since $\{f_n(x)\}$ is monotone increasing in n, the same is true of $\{\int_E f_n\,d\mu\}$. Therefore $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{R}^*$. Let us call this limit k. For each n, $\int_E f_n\,d\mu \le \int_E f\,d\mu$ because $f_n \le f$. Therefore $k \le \int_E f\,d\mu$, and the problem is to prove the reverse inequality.

Let c be any real number with $0 < c < 1$, to be regarded as close to 1, and let s be a simple function with $0 \le s \le f$. Define

$$E_n = \{x \in E \mid f_n(x) \ge cs(x)\}.$$

These sets are measurable, and $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \subseteq E$. Let us see that $E = \bigcup_{n=1}^\infty E_n$. If $f(x) = 0$ for a particular x in E, then $f_n(x) = 0$ for all n and also $cs(x) = 0$. Thus x is in every E_n. If $f(x) > 0$, then the inequality $f(x) \ge s(x)$ forces $f(x) > cs(x)$. Since $f_n(x)$ has increasing limit $f(x)$, $f_n(x)$ must be $> cs(x)$ eventually, and then x is in E_n. In either case x is in $\bigcup_{n=1}^\infty E_n$. Thus $E = \bigcup_{n=1}^\infty E_n$.

For every n, we have

$$k \ge \int_E f_n\,d\mu \ge \int_{E_n} f_n\,d\mu \ge \int_{E_n} cs\,d\mu = c\int_{E_n} s\,d\mu.$$

Since, by Theorem 5.19, the integral is a completely additive set function, Proposition 5.2 shows that $\lim \int_{E_n} s\,d\mu = \int_E s\,d\mu$. Therefore $k \ge c\int_E s\,d\mu$. Since c is arbitrary with $0 < c < 1$, $k \ge \int_E s\,d\mu$. Taking the supremum over s with $0 \le s \le f$, we conclude that $k \ge \int_E f\,d\mu$. $\square$

Proposition 5.26. If f and g are measurable functions, if their sum $h = f + g$ is everywhere defined, and if $\int_E f\,d\mu + \int_E g\,d\mu$ is defined, then $\int_E h\,d\mu$ is defined and

$$\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu.$$

REMARK. It may seem surprising that complete additivity plays a role in the proof of this proposition, since it apparently played no role in the linearity of the Riemann integral. In fact, although complete additivity is used when f and g are unbounded, it can be avoided when f and g are bounded, as will be observed in Problems 42–43 at the end of the chapter. The distinction between the two cases is that the pointwise convergence in Proposition 5.11 is actually uniform if the given function is bounded, whereas it cannot be uniform for an unbounded function because the uniform limit of bounded functions is bounded.

PROOF. The sum h is measurable by Proposition 5.7. For the conclusions about integration, first assume that $f \geq 0$ and $g \geq 0$. In the case of simple functions $s = t + u$ with $t \geq 0$ and $u \geq 0$, we use Proposition 5.14 and Lemma 5.12. The proposition shows that we are to prove that $\mathcal{I}_E(s) = \mathcal{I}_E(t) + \mathcal{I}_E(u)$, and the lemma shows that we can use expansions of t and u into sets on which t and u are both constant and the conclusion about $\mathcal{I}_E(s)$ is evident. If f and g are ≥ 0 and are not necessarily simple, then we can use Proposition 5.11 to find increasing sequences $\{t_n\}$ and $\{u_n\}$ of simple functions ≥ 0 with limits f and g. If $s_n = t_n + u_n$, then s_n is nonnegative simple, and $\{s_n\}$ increases to h. For each n, we have just proved that $\int_E s_n\,d\mu = \int_E t_n\,d\mu + \int_E u_n\,d\mu$, and therefore $\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$ by the Monotone Convergence Theorem (Theorem 5.25).

The next case is that $f \geq 0$, $g \leq 0$, and $h = f + g \geq 0$. Then $f = h + (-g)$ with $h \geq 0$ and $(-g) \geq 0$, so that $\int_E f\,d\mu = \int_E h\,d\mu + \int_E (-g)\,d\mu$. Hence $\int_E h\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$, provided the right side is defined.

For a general $h \geq 0$, we decompose E into the disjoint union of three sets, one where $f \geq 0$ and $g \geq 0$, one where $f \geq 0$ and $g < 0$, and one where $f < 0$ and $g \geq 0$. The additivity of the integral as a set function (Theorem 5.19), in combination with the cases that we have already proved, then gives the desired result. Finally for general h, we have only to write $h = h^+ - h^-$ and consider h^+ and h^- separately. □

Corollary 5.27. Let E be a measurable set, and let $\{f_n\}$ be a sequence of measurable functions ≥ 0. Put $F(x) = \sum_{n=1}^{\infty} f_n(x)$. Then $\int_E F\,d\mu = \sum_{n=1}^{\infty} \int_E f_n\,d\mu$.

PROOF. Apply Proposition 5.26 to the n^{th} partial sum of the series, and then use the Monotone Convergence Theorem (Theorem 5.25). □

The next corollary is given partly to illustrate a standard technique for passing from integration results about indicator functions to integration results about general functions. This technique is used again and again in measure theory.

Corollary 5.28. If $f \geq 0$ is a measurable function and if ν is the measure $\nu(E) = \int_E f\,d\mu$, then $\int_E g\,d\nu = \int_E gf\,d\mu$ for every measurable function g for which at least one side is defined.

REMARKS. The set function ν is a measure by Theorem 5.19. In the situation of this corollary, we shall write $\nu = f\ d\mu$.

PROOF. By Corollary 5.20 it is enough to prove that

$$\int_X g\,d\nu = \int_X gf\,d\mu. \tag{$*$}$$

For $g = I_E$, $(*)$ is true by hypothesis. Proposition 5.26 shows that $(*)$ extends to be valid for simple functions $g \geq 0$. For general $g \geq 0$, Proposition 5.11 produces an increasing sequence $\{s_n\}$ of simple functions ≥ 0 with pointwise limit g. Then $(*)$ for this g follows from the result for simple functions in combination with monotone convergence. For general g, write $g = g^+ - g^-$, apply $(*)$ for g^+ and g^-, and subtract the results using Proposition 5.26. $\square$

Theorem 5.29 (Fatou's Lemma). If E is a measurable set and if $\{f_n\}$ is a sequence of *nonnegative* measurable functions, then

$$\int_E \liminf_n f_n\,d\mu \leq \liminf_n \int_E f_n\,d\mu.$$

In particular, if $f(x) = \lim_n f_n(x)$ exists for all x, then

$$\int_E f\,d\mu \leq \liminf_n \int_E f_n\,d\mu.$$

REMARK. Fatou's Lemma applies to both examples that precede the Monotone Convergence Theorem (Theorem 5.25), and strict inequality holds in both cases.

PROOF. Set $g_n(x) = \inf_{k\geq n} f_k(x)$. Then $\lim_n g_n(x) = \liminf f_n(x)$, and the Monotone Convergence Theorem (Theorem 5.25) gives

$$\int_E \liminf_n f_n\,d\mu = \int_E \lim_n g_n\,d\mu = \lim_n \int_E g_n\,d\mu.$$

But $g_n(x) \leq f_n(x)$ pointwise, so that $\int_E g_n\,d\mu \leq \int_E f_n\,d\mu$ for all n. Thus

$$\lim \int_E g_n\,d\mu \leq \liminf \int_E f_n\,d\mu,$$

and the theorem follows. $\square$

Theorem 5.30 (Dominated Convergence Theorem). Let E be a measurable set, and suppose that $\{f_n\}$ is a sequence of measurable functions such that for some *integrable* g, $|f_n| \leq g$ for all n. If $f = \lim f_n$ exists pointwise, then $\lim_n \int_E f_n \, d\mu$ exists, f is integrable on E, and

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

PROOF. The set on which g is infinite has measure 0, since g is integrable. If we redefine g, f_n, and f to be 0 on this set, we change no integrals and we affect the validity of neither the hypotheses nor the conclusion.

By Corollary 5.18, f is integrable on E, and so is f_n for every n. Applying Fatou's Lemma (Theorem 5.29) to $f_n + g \geq 0$, we obtain $\int_E (f + g) \, d\mu \leq \liminf \int_E (f_n + g) \, d\mu$. Since g is integrable and everywhere finite, this inequality becomes

$$\int_E f \, d\mu \leq \liminf \int_E f_n \, d\mu.$$

A second application of Fatou's Lemma, this time to $g - f_n \geq 0$, gives $\int_E (g - f) \, d\mu \leq \liminf \int_E (g - f_n) \, d\mu$. Thus

$$-\int_E f \, d\mu \leq \liminf \int_E (-f_n) \, d\mu$$

and

$$\int_E f \, d\mu \geq \limsup \int_E f_n \, d\mu.$$

Therefore $\lim \int_E f_n \, d\mu$ exists and has the value asserted. $\square$

Corollary 5.31. Let E be a set of finite measure, let $c \geq 0$ be in $\mathbb{R}$, and suppose that $\{f_n\}$ is a sequence of measurable functions such that $|f_n| \leq c$ for all n. If $f = \lim f_n$ exists pointwise, then $\lim \int_E f_n \, d\mu$ exists, f is integrable on E, and

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

PROOF. This is the special case $g = c$ in Theorem 5.30. $\square$

5. Proof of the Extension Theorem

In this section we shall prove the Extension Theorem, Theorem 5.5. After the end of the proof, we shall fill in one further detail left from Section 1—to show that a measure on a σ-ring has a canonical extension to a measure on the smallest σ-algebra containing the given σ-ring.

Most of this section will concern the proof of the Extension Theorem in the case that X is measurable and $\nu(X)$ is finite. Thus, until further notice, let us assume that X is a nonempty set, $\mathcal{A}$ is an algebra of subsets of X, and ν is a nonnegative completely additive set function defined on $\mathcal{A}$ such that $\nu(X) < +\infty$.

In a way, the intuition for the proof is typical of that for many existence-uniqueness theorems in mathematics: to see how to prove existence, we assume existence and uniqueness outright, see what necessary conditions each of the assumptions puts on the object to be constructed, and then begin the proof.

With the present theorem in the case that $\nu(X)$ is finite, we shall assign to each subset E of X an upper bound $\mu^*(E)$ and a lower bound $\mu_*(E)$ for the value of the extended measure on the set E. If the existence half of the theorem is valid, we must have $\mu_*(E) \leq \mu^*(E)$ for E in the smallest σ-algebra containing $\mathcal{A}$. In fact, we shall see that this inequality holds for *all* subsets E of X. On the other hand, if $\mu_*(E) < \mu^*(E)$ for some E in the σ-algebra of interest and if our upper and lower bounds are good estimates, we might expect that there is more than one way to define the extended measure on E, in contradiction to uniqueness. That thought suggests trying to prove that $\mu_*(E) = \mu^*(E)$ for the sets of interest. One way of doing so is to try to prove that the class of subsets for which this equality holds is a σ-algebra containing $\mathcal{A}$, and then the common value of μ_* and μ^* is the desired extension.

This procedure in fact works, and the only subtlety is in the definitions of $\mu_*(E)$ and $\mu^*(E)$. We give these definitions after one preliminary lemma that will make μ_* and μ^* well defined. For orientation, think of the setting as the unit interval [0, 1], with Lebesgue measure to be extended from the elementary sets to the Borel sets. In this case the families $\mathcal{U}$ and $\mathcal{K}$ in the first lemma contain all the open sets and all the compact sets, respectively, and may be regarded as generalizations of these collections of sets.

Lemma 5.32. Let $\mathcal{U}$ be the class of all countable unions of sets in $\mathcal{A}$, and let $\mathcal{K}$ be the class of all countable intersections of sets in $\mathcal{A}$. Then μ^* and μ_* are consistently defined on $\mathcal{U}$ and $\mathcal{K}$, respectively, by letting

$$\mu^*(U) = \lim \nu(A_n) \qquad \text{and} \qquad \mu_*(K) = \lim \nu(C_n)$$

whenever $\{A_n\}$ is an increasing sequence of sets in $\mathcal{A}$ with union U and $\{C_n\}$ is a decreasing sequence of sets in $\mathcal{A}$ with intersection K. Moreover, μ^* and μ_* have the following properties:

(a) μ^* and μ_* agree with ν on sets of $\mathcal{A}$,
(b) $\mu^*(U) \leq \mu^*(V)$ whenever U is in $\mathcal{U}$, V is in $\mathcal{U}$, and $U \subseteq V$,

(c) $\mu_*(K) \le \mu_*(L)$ whenever K is in $\mathcal{K}$, L is in $\mathcal{K}$, and $K \subseteq L$,
(d) $\lim \mu^*(U_n) = \mu^*(U)$ whenever $\{U_n\}$ is an increasing sequence of sets in $\mathcal{U}$ with union U.

PROOF. If $\{B_n\}$ is another increasing sequence in $\mathcal{A}$ with union U, then Proposition 5.2 and Theorem 1.13 give

$$\lim_m \nu(A_m) = \lim_m \big(\lim_n \nu(A_m \cap B_n)\big) = \lim_n \big(\lim_m \nu(A_m \cap B_n)\big) = \lim_n \nu(B_n).$$

Hence μ^* is consistently defined on $\mathcal{U}$. Similarly if $\{D_n\}$ decreases to K, then Corollary 5.3 and Theorem 1.13 give

$$\begin{aligned}\nu(X) - \lim_m \nu(C_m) &= \nu(X) - \lim_m \big(\lim_n \nu(C_m \cap D_n)\big)\\ &= \nu(X) - \lim_n \big(\lim_m \nu(C_m \cap D_n)\big) = \nu(X) - \lim_n \nu(D_n),\end{aligned}$$

and hence $\lim_m \nu(C_m) = \lim_n \nu(D_n)$. Thus μ_* is consistently defined on $\mathcal{K}$. The set functions μ^* and μ_* are defined on all of $\mathcal{U}$ and $\mathcal{K}$ because a set that is a countable union (or intersection) of sets in an algebra is a countable increasing union (or decreasing intersection).

Of the four properties, (a) is clear, and (b) and (c) follow from the inequalities

$$\mu^*(U) = \sup_{A \subseteq U,\, A \in \mathcal{A}} \nu(A) \le \sup_{A \subseteq V,\, A \in \mathcal{A}} \nu(A) = \mu^*(V)$$

and

$$\mu_*(K) = \inf_{A \supseteq K,\, A \in \mathcal{A}} \nu(A) \le \inf_{A \supseteq L,\, A \in \mathcal{A}} \nu(A) = \mu_*(L).$$

In (d), U is in $\mathcal{U}$, since the countable union of countable unions is again a countable union, and (b) shows that $\lim \mu^*(U_n) \le \mu^*(U)$. For each n, let $\{A_m^{(n)}\}$ be an increasing sequence of sets from $\mathcal{A}$ with union U_n. Arrange all the $A_m^{(n)}$ in a sequence, and let B_k denote the union of the first k members of the sequence. Then $\{B_k\}$ is an increasing sequence with union U. Let $\epsilon > 0$ be given, and choose M large enough so that $\mu^*(B_M) \ge \mu^*(U) - \epsilon$. Since the sets U_n increase, since B_M is a finite union of sets $A_m^{(n)}$, and since $A_m^{(n)} \subseteq U_n$, we must have $\mu^*(U_N) \ge \mu^*(B_M)$ for some N. But then

$$\lim \mu^*(U_n) \ge \mu^*(U_N) \ge \mu^*(B_M) \ge \mu^*(U) - \epsilon.$$

Since ϵ is arbitrary, $\lim \mu^*(U_n) \ge \mu^*(U)$. □

For each subset E of X, we define

$$\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U) \qquad \text{and} \qquad \mu_*(E) = \sup_{K \subseteq E,\, K \in \mathcal{K}} \mu_*(K).$$

Conclusions (b) and (c) of Lemma 5.32 show that the new definitions of μ^* and μ_* are consistent with the old ones. The set functions μ^* and μ_* on arbitrary subsets E of X may be called the **outer measure** and the **inner measure** associated to ν.

Lemma 5.33. If A and B are subsets of X with $A \subseteq B$, then $\mu^*(A) \le \mu^*(B)$ and $\mu_*(A) \le \mu_*(B)$. In addition,

(a) if $E \subseteq \bigcup_{n=1}^\infty E_n$, then $\mu^*(E) \le \sum_{n=1}^\infty \mu^*(E_n)$,
(b) if F and G are disjoint, then $\mu_*(F) + \mu_*(G) \le \mu_*(F \cup G)$.

PROOF. Since $\mu^*(A)$ is an infimum over a larger class of sets than $\mu^*(B)$ is, we have $\mu^*(A) \le \mu^*(B)$. Similarly $\mu_*(A) \le \mu_*(B)$.

For (a), let $E \subseteq \bigcup_{n=1}^\infty E_n$. In the special case in which E_n is in $\mathcal{U}$ for all n, let $\{F_m^{(n)}\}$ be, for fixed n and varying m, an increasing sequence of sets in $\mathcal{A}$ with union E_n. For any N, we then have $\bigcup_{m=1}^\infty (F_m^{(1)} \cup \cdots \cup F_m^{(N)}) = E_1 \cup \cdots \cup E_N$. Hence

$$\begin{aligned}
\mu^*(E) \le \mu^*\Big(\bigcup_{n=1}^\infty E_n\Big) &= \lim_N \mu^*\Big(\bigcup_{n=1}^N E_n\Big) && \text{by Lemma 5.32d} \\
&= \lim_N \mu^*\Big(\bigcup_{m=1}^\infty (F_m^{(1)} \cup \cdots \cup F_m^{(N)})\Big) \\
&= \lim_N \lim_m \nu(F_m^{(1)} \cup \cdots \cup F_m^{(N)}) && \text{by definition of } \mu^* \text{ on } \mathcal{U} \\
&\le \lim_N \lim_m \sum_{n=1}^N \nu(F_m^{(n)}) && \text{by Proposition 5.1f} \\
&= \lim_N \sum_{n=1}^N \mu^*(E_n) = \sum_{n=1}^\infty \mu^*(E_n).
\end{aligned}$$

For general subsets E_n of X, choose U_n in $\mathcal{U}$ with $U_n \supseteq E_n$ and $\mu^*(U_n) \le \mu^*(E_n) + \epsilon/2^n$. Then $E \subseteq \bigcup_n U_n$, and the special case applied to the U_n shows that

$$\mu^*(E) \le \mu^*\Big(\bigcup_n U_n\Big) \le \sum_n \mu^*(U_n) \le \sum_n \mu^*(E_n) + \epsilon.$$

Hence $\mu^*(E) \le \sum_n \mu^*(E_n)$, and (a) is proved.

For (b), let F and G be disjoint. In the special case in which F and G are in $\mathcal{K}$, let $\{F_n\}$ and $\{G_n\}$ be decreasing sequences of sets in $\mathcal{A}$ with intersections F and G. Then

$$\begin{aligned}
\mu_*(F \cup G) &= \lim \nu(F_n \cup G_n) && \text{by definition of } \mu_* \text{ on } \mathcal{K} \\
&= \lim \big(\nu(F_n) + \nu(G_n) - \nu(F_n \cap G_n)\big) && \text{by Proposition 5.1b} \\
&= \mu_*(F) + \mu_*(G),
\end{aligned}$$

the last step holding by Corollary 5.3, since $F \cap G$ is empty. For general disjoint subsets F and G in X, choose K and L in $\mathcal{K}$ with $K \subseteq F$, $L \subseteq G$, $\mu_*(K) \geq \mu_*(F) - \epsilon$, and $\mu_*(L) \geq \mu_*(G) - \epsilon$. Then

$$\mu_*(F \cup G) \geq \mu_*(K \cup L) = \mu_*(K) + \mu_*(L) \geq \mu_*(F) + \mu_*(G) - 2\epsilon,$$

the middle step holding by the special case. Hence $\mu_*(F \cup G) \geq \mu_*(F) + \mu_*(G)$, and (b) is proved. $\square$

Lemma 5.34. For every subset E of X, $\mu_*(E) \leq \mu^*(E)$. Equality holds if E is in $\mathcal{U}$ or $\mathcal{K}$.

PROOF. The proof is in three steps.

First we prove that if U is in $\mathcal{U}$ and K is in $\mathcal{K}$, then $\mu^*(U) \leq \mu_*(U)$ and $\mu^*(K) \leq \mu_*(K)$. In fact, choose C in $\mathcal{A}$ with $C \subseteq U$ and $\mu^*(U) \leq \nu(C) + \epsilon$. Then $\mu^*(U) \leq \nu(C) + \epsilon \leq \mu_*(U) + \epsilon$ by Lemma 5.33 since $C \subseteq U$. Hence $\mu^*(U) \leq \mu_*(U)$. Similarly choose D in $\mathcal{A}$ with $D \supseteq K$ and $\mu_*(K) \geq \nu(D) - \epsilon$. Then $\mu_*(K) \geq \nu(D) - \epsilon \geq \mu^*(K) - \epsilon$, and hence $\mu_*(K) \geq \mu^*(K)$.

Second we prove that if K is in $\mathcal{K}$, then $\mu^*(K) = \mu_*(K)$. In fact, choose C in $\mathcal{A}$ with $C \supseteq K$ and $\nu(C) - \mu_*(K) \leq \epsilon$. Then $C - K$ is in $\mathcal{U}$, and

$$\begin{aligned}
\mu_*(K) &\leq \nu(C) \leq \mu^*(C - K) + \mu^*(K) && \text{by Lemma 5.33a} \\
&\leq \big(\mu_*(C - K) + \mu_*(K)\big) - \mu_*(K) + \mu^*(K) && \text{by the previous step} \\
&\leq \nu(C) - \mu_*(K) + \mu^*(K) && \text{by Lemma 5.33b} \\
&\leq \mu^*(K) + \epsilon && \text{by the choice of } C.
\end{aligned}$$

Combining this inequality with the previous step, we see that $\mu^*(K) = \mu_*(K)$.

Third we prove that $\mu_*(E) \leq \mu^*(E)$ for every E. In fact, find K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$, $\mu_*(K) \geq \mu_*(E) - \epsilon$, and $\mu^*(U) \leq \mu^*(E) + \epsilon$. Then $\mu_*(E) \leq \mu_*(K) + \epsilon = \mu^*(K) + \epsilon \leq \mu^*(U) + \epsilon \leq \mu^*(E) + 2\epsilon$, and the proof is complete. $\square$

Define a subset E of X to be **measurable** for purposes of this section if $\mu_*(E) = \mu^*(E)$, and let $\mathcal{B}$ be the class of measurable subsets of X. Lemma 5.34 shows that $\mathcal{U}$ and $\mathcal{K}$ are both contained in $\mathcal{B}$.

Lemma 5.35. If U is in $\mathcal{U}$ and K is in $\mathcal{K}$ with $K \subseteq U$, then

$$\mu^*(U - K) = \mu^*(U) - \mu_*(K).$$

If E is measurable, then for any $\epsilon > 0$, there are sets K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$ and

$$\mu^*(E - K) \leq \mu^*(U - K) \leq \epsilon.$$

PROOF. For the first conclusion, $U - K$ is in $\mathcal{U}$ and hence $\mu^*(U - K) = \mu_*(U - K) = \mu_*(U) - \mu_*(K) = \mu^*(U) - \mu_*(K)$ by Lemma 5.34, Lemma 5.33b, and Lemma 5.34 again.

For the second conclusion choose K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$, $\mu_*(K) + \frac{\epsilon}{2} \geq \mu_*(E)$, and $\mu^*(E) \geq \mu^*(U) - \frac{\epsilon}{2}$. Since $\mu_*(E) = \mu^*(E)$ by the assumed measurability, we see that $\mu_*(K) + \frac{\epsilon}{2} \geq \mu^*(U) - \frac{\epsilon}{2}$, hence that $\mu^*(U) - \mu_*(K) \leq \epsilon$. The result now follows from Lemma 5.33 and the first conclusion of the present lemma. $\square$

Lemma 5.36. The class $\mathcal{B}$ of measurable sets is a σ-algebra containing $\mathcal{A}$, and the restriction of μ^* to $\mathcal{B}$ is a measure.

PROOF. Certainly $\mathcal{B} \supseteq \mathcal{A}$. The rest of the proof is in three steps.

First we prove that the intersection of two measurable sets is measurable. In fact, let F and G be in $\mathcal{B}$, and use Lemma 5.35 to choose $K \subseteq F$ and $L \subseteq G$ with $\mu^*(F-K) \leq \epsilon$ and $\mu^*(G-L) \leq \epsilon$. Since $F \cap G \subseteq (F-K) \cup (K \cap L) \cup (G-L)$,

$$\begin{aligned}
&\mu^*(F \cap G) \\
&\qquad \leq \mu^*(F - K) + \mu^*(K \cap L) + \mu^*(G - L) && \text{by Lemma 5.33a} \\
&\qquad \leq \mu^*(K \cap L) + 2\epsilon && \text{by definition of } K \text{ and } L \\
&\qquad = \mu_*(K \cap L) + 2\epsilon && \text{by Lemma 5.34} \\
&\qquad \leq \mu_*(F \cap G) + 2\epsilon && \text{since } K \cap L \subseteq F \cap G.
\end{aligned}$$

Second we prove that the complement of a measurable set is measurable. Let E be measurable. By Lemma 5.35 choose K in $\mathcal{K}$ and U in $\mathcal{U}$ with $K \subseteq E \subseteq U$ and $\mu^*(U - K) \leq \epsilon$. Since $U^c \subseteq E^c \subseteq K^c$ and $K^c - U^c = U - K$, we have

$$\begin{aligned}
\mu^*(E^c) &\leq \mu^*(K^c - U^c) + \mu^*(U^c) && \text{by Lemma 5.33a} \\
&= \mu^*(U - K) + \mu_*(U^c) && \text{since } U^c \text{ is in } \mathcal{K} \\
&\leq \epsilon + \mu_*(E^c).
\end{aligned}$$

Thus the complement of a measurable set is measurable, and $\mathcal{B}$ is an algebra of sets.

Third we prove that the countable disjoint union of measurable sets is measurable, and μ^* is a measure on $\mathcal{B}$. In fact, let $\{E_n\}$ be a sequence of disjoint sets in $\mathcal{B}$. Application of Lemma 5.33a, Lemma 5.33b, and Lemma 5.34 gives

$$\begin{aligned}
\mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big) &\leq \sum_{n=1}^{\infty} \mu^*(E_n) = \sum_{n=1}^{\infty} \mu_*(E_n) = \lim_N \sum_{n=1}^{N} \mu_*(E_n) \\
&\leq \lim_N \mu_*\Big(\bigcup_{n=1}^{N} E_n\Big) \leq \mu_*\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leq \mu^*\Big(\bigcup_{n=1}^{\infty} E_n\Big).
\end{aligned}$$

The end members of this chain of inequalities are equal, and thus equality must hold throughout: $\mu_*(\bigcup_n E_n) = \mu^*(\bigcup_n E_n) = \sum \mu^*(E_n)$. Consequently $\bigcup_n E_n$ is measurable, and μ^* is completely additive. $\square$

PROOF OF THEOREM 5.5 UNDER THE SPECIAL HYPOTHESES. We continue to assume that the given ring of subsets of X is an algebra and that $\nu(X)$ is finite. Define $\mathcal{B}$ to be the class of measurable sets in the previous construction. Then Lemma 5.36 shows that $\mathcal{B}$ is a σ-algebra containing $\mathcal{A}$. Hence $\mathcal{B}$ contains the smallest σ-algebra $\mathcal{C}$ containing $\mathcal{A}$. Lemma 5.36 shows also that the restriction of μ^* to $\mathcal{C}$ is a measure extending ν. This proves existence of the extension under the special hypotheses.

For uniqueness, suppose that μ' is an extension of ν to $\mathcal{C}$. Proposition 5.2 and Corollary 5.3 show that μ' has to agree with μ^* on $\mathcal{U}$ and with μ_* on $\mathcal{K}$. If $K \subseteq E \subseteq U$ with K in $\mathcal{K}$ and U in $\mathcal{U}$, then we have

$$\mu_*(K) = \mu'(K) \leq \mu'(E) \leq \mu'(U) = \mu^*(U).$$

Taking the supremum over K and the infimum over U gives $\mu_*(E) \leq \mu'(E) \leq \mu^*(E)$. Since E is in $\mathcal{B}$, $\mu_*(E) = \mu^*(E)$, and we see that $\mu'(E) = \mu^*(E)$. Thus μ' coincides with the restriction of μ^* to $\mathcal{C}$. This proves uniqueness of the extension under the special hypotheses. $\square$

Now we return to the general hypotheses of Theorem 5.5—that $\mathcal{R}$ is a ring of subsets of X, that ν is a nonnegative completely additive set function on $\mathcal{R}$, and that ν is σ-finite—and we shall complete the proof that ν extends uniquely to a measure on the smallest σ-ring $\mathcal{C}$ containing $\mathcal{R}$.

PROOF OF THEOREM 5.5 IN THE GENERAL CASE. If S is an element of $\mathcal{R}$ with $\nu(S)$ finite, define $S \cap \mathcal{R} = \{S \cap R \mid R \in \mathcal{R}\}$. Then $(S, S \cap \mathcal{R}, \nu\big|_{S\cap\mathcal{R}})$ is a set of data satisfying the special hypotheses of the Extension Theorem considered above. By the special case, if $\mathcal{C}_S$ denotes the smallest σ-algebra of subsets of S containing $S \cap \mathcal{R}$, then $\nu\big|_{S\cap\mathcal{R}}$ has a unique extension to a measure μ_S on $\mathcal{C}_S$. The measures μ_S have a certain consistency property because of the uniqueness: if $S' \subseteq S$, then $\mu_S\big|_{S'\cap\mathcal{R}} = \mu_{S'}$.

Now let $\{S_n\}$ be a sequence of sets in $\mathcal{R}$ with union S in $\mathcal{C}$ and with $\nu(S_n)$ finite for all n. Possibly replacing each set S_n by the difference of S_n and all previous S_k's, we may assume that the sequence is disjoint. We define μ_S on the σ-algebra $S \cap \mathcal{C}$ of subsets S by $\mu_S(E) = \sum_n \mu_{S_n}(E \cap S_n)$ for E in $S \cap \mathcal{C}$. Let us check that μ_S is unambiguously defined and is completely additive. If $\{T_m\}$ is another sequence of sets in $\mathcal{R}$ with union S and with $\nu(T_m)$ finite for all m, then the corresponding definition of a set function on $S \cap \mathcal{C}$ is $\mu'_S(E) = \sum_m \mu_{T_m}(E \cap T_m)$. The consistency property from the previous paragraph gives us $\mu_{S_n}(E \cap S_n \cap T_m) = \mu_{T_m}(E \cap S_n \cap T_m)$. Then Corollary 1.15 allows us to write

$$\begin{aligned}\mu'_S(E) &= \sum_m \mu_{T_m}(E \cap T_m) = \sum_m \sum_n \mu_{T_m}(E \cap S_n \cap T_m) \\ &= \sum_m \sum_n \mu_{S_n}(E \cap S_n \cap T_m) = \sum_n \sum_m \mu_{S_n}(E \cap S_n \cap T_m) \\ &= \sum_n \mu_{S_n}(E \cap S_n) = \mu_S(E),\end{aligned}$$

and we see that μ_S is unambiguously defined. To check that μ_S is completely additive, let $F_1, F_2, \ldots$ be a disjoint sequence of sets in $S \cap \mathcal{C}$ with union F. Then the complete additivity of μ_{S_n}, in combination with Corollary 1.15, gives

$$\begin{aligned}\mu_S(F) &= \sum_n \mu_{S_n}(F \cap S_n) = \sum_n \sum_m \mu_{S_n}(F_m \cap S_n) \\ &= \sum_m \sum_n \mu_{S_n}(F_m \cap S_n) = \sum_m \mu_S(F_m),\end{aligned}$$

and thus μ_S is completely additive.

The measures μ_S are consistent on their common domains. To see the consistency, let us see that μ_S and μ_T agree on subsets of $S \cap T$. Let S be the countable disjoint union of sets S_n in $\mathcal{R}$, and let T be the countable disjoint union of sets T_m in $\mathcal{R}$. Then $S \cap T$ is the countable disjoint union of the sets $S_n \cap T_m$. If E is in $(S \cap T) \cap \mathcal{C}$, then Corollary 1.15 and the consistency property of the set functions μ_R for R in $\mathcal{R}$ yield

$$\begin{aligned}\mu_S(E) &= \sum_n \mu_{S_n}(E \cap S_n) = \sum_n \mu_{S_n}(E \cap S_n \cap T) \\ &= \sum_n \sum_m \mu_{S_n}(E \cap S_n \cap T_m) = \sum_n \sum_m \mu_{S_n \cap T_m}(E \cap S_n \cap T_m) \\ &= \sum_m \sum_n \mu_{S_n \cap T_m}(E \cap S_n \cap T_m) = \sum_m \sum_n \mu_{T_m}(E \cap S_n \cap T_m) \\ &= \sum_m \mu_{T_m}(E \cap S \cap T_m) = \sum_m \mu_{T_m}(E \cap T_m) = \mu_T(E).\end{aligned}$$

Hence the measures μ_S are consistent on their common domains.

If $\mathcal{M}$ denotes the set of subsets of X that are contained in a countable union of members of $\mathcal{R}$ on which ν is finite, then $\mathcal{M}$ is closed under countable unions and differences and is thus a σ-ring containing $\mathcal{R}$. It therefore contains $\mathcal{C}$, and we conclude that every member of $\mathcal{C}$ is contained in a countable union of members of $\mathcal{R}$ on which ν is finite. It follows that we can define μ on all of $\mathcal{C}$ as follows: if E

is in $\mathcal{C}$, then E is contained in some countable union S of members of $\mathcal{R}$ on which ν is finite, and we define $\mu(E) = \mu_S(E)$. We have seen that the measures μ_S are consistently defined, and hence $\mu(E)$ is well defined. If a countable disjoint union $E = \bigcup_{N=1}^{\infty} E_n$ of sets in $\mathcal{C}$ is given, then all the sets in question lie in a single S, and we then have $\mu(E) = \mu_S(E) = \sum_{n=1}^{\infty} \mu_S(E_n) = \sum_{n=1}^{\infty} \mu(E_n)$. In other words, μ is completely additive. This proves existence.

For uniqueness let E be given in $\mathcal{C}$, and suppose that S is a member of $\mathcal{C}$ containing E and equal to the countable disjoint union of sets S_n in $\mathcal{R}$ with $\nu(S_n)$ finite for all n. We have seen that the value of $\mu(E \cap S_n) = \mu_{S_n}(E \cap S_n)$ is determined by $\nu\big|_{S_n \cap \mathcal{R}}$, hence by ν on $\mathcal{R}$. By complete additivity of μ, $\mu(E)$ is determined by the values of $\mu(E \cap S_n)$ for all n. Therefore μ on $\mathcal{C}$ is determined by ν on $\mathcal{R}$. This proves uniqueness. $\square$

As was promised, we shall now fill in one further detail left from Section 1—to show that a measure on a σ-ring has a canonical extension to a measure on the smallest σ-algebra containing the given σ-ring.

Proposition 5.37. Let $\mathcal{R}$ be a σ-ring of subsets of a nonempty set X, let $\mathcal{R}_c$ be the set of complements in X of the members of $\mathcal{R}$, and let $\mathcal{A}$ be the smallest σ-algebra containing $\mathcal{R}$. Then either

(i) $\mathcal{R} = \mathcal{R}_c = \mathcal{A}$ or
(ii) $\mathcal{R} \cap \mathcal{R}_c = \varnothing$ and $\mathcal{A} = \mathcal{R} \cup \mathcal{R}_c$.

In the latter case any measure μ on $\mathcal{R}$ has a canonical extension to a measure μ_1 on $\mathcal{A}$ given by $\mu_1(E) = \sup\big\{\mu(F) \bigm| F \in \mathcal{R} \text{ and } F \subseteq E\big\}$ for E in $\mathcal{R}_c$. This canonical extension has the property that any other extension μ_2 satisfies $\mu_2 \geq \mu_1$.

PROOF. If X is in $\mathcal{R}$, then $\mathcal{R}$ is closed under complements, since $\mathcal{R}$ is closed under differences; hence $\mathcal{R} = \mathcal{R}_c = \mathcal{A}$. If X is not in $\mathcal{R}$, then $\mathcal{R} \cap \mathcal{R}_c = \varnothing$ because any set E in the intersection has E^c in the intersection and then also $X = E \cup E^c$ in the intersection. In this latter case it is plain that $\mathcal{A} \supseteq \mathcal{R} \cup \mathcal{R}_c$. Thus (ii) will be the only alternative to (i) if it is proved that $\mathcal{B} = \mathcal{R} \cup \mathcal{R}_c$ is a σ-algebra. Certainly $\mathcal{B}$ is closed under complements. To see that $\mathcal{B}$ is closed under countable unions, we may assume, because $\mathcal{R}$ is a σ-ring, that we are to check the union of countably many sets with at least one in $\mathcal{R}_c$. Thus let $\{E_n\}$ be a sequence of sets in $\mathcal{R}$, and let $\{F_n\}$ be a sequence of sets in $\mathcal{R}_c$. Then $E = \bigcup_{n=1}^{\infty} E_n$ is in $\mathcal{R}$ and $F = \bigcap_{n=1}^{\infty} F_n^c$ is in $\mathcal{R}$, since $\mathcal{R}$ is a σ-ring. The union of the sets E_n and F_n in question is $E \cup F^c = (F - E)^c$, is exhibited as the complement of the difference of two sets in $\mathcal{R}$, and is therefore in $\mathcal{R}_c$. Thus $\mathcal{A}$ is closed under countable unions and is a σ-algebra.

In the case of (ii), let us see that μ_1 is a measure on $\mathcal{A}$. If we are to check the measure of a disjoint sequence of sets in $\mathcal{A}$, there is no problem if all the sets are in $\mathcal{R}$, since $\mu_1\big|_{\mathcal{R}} = \mu$ is completely additive. There cannot be as many as two of the sets in $\mathcal{R}_c$ because no two sets F_1 and F_2 in $\mathcal{R}_c$ are disjoint; in fact, $F_1 \cap F_2 = (F_1^c \cup F_2^c)^c$ exhibits the intersection as in $\mathcal{R}_c$, and the empty set is not a member of $\mathcal{R}_c$. Thus we may assume that the disjoint sequence consists of a sequence $\{E_n\}$ of sets in $\mathcal{R}$ and a single set F in $\mathcal{R}_c$. If $E = \bigcup_{n=1}^\infty E_n$, then $\mu_1(E) = \mu(E) = \sum_{n=1}^\infty \mu(E_n) = \sum_{n=1}^\infty \mu_1(E_n)$. So it is enough to see that $\mu_1(E \cup F) = \mu(E) + \mu_1(F)$. If E' is a subset of F that is in $\mathcal{R}$, then $\mu_1(E \cup F) \geq \mu(E \cup E') = \mu(E) + \mu(E')$. Taking the supremum over all such E' shows that $\mu_1(E \cup F) \geq \mu(E) + \mu_1(F)$. For the reverse inequality let S be a member of $\mathcal{R}$ contained in $E \cup F$. Then the sets $E \cap S$ and $F \cap S = S - F^c$ are in $\mathcal{R}$, and thus $\mu(S) = \mu(E \cap S) + \mu(F \cap S) \leq \mu(E) + \mu_1(F)$. Taking the supremum over S gives $\mu_1(E \cup F) \leq \mu(E) + \mu_1(F)$. Thus μ_1 is completely additive.

If μ_2 is any other extension, any set F in $\mathcal{R}^c$ has $\mu_2(F) \geq \mu_2(E) = \mu(E)$ for all subsets E of F that are in $\mathcal{R}$. Taking the supremum over E gives $\mu_2(F) \geq \mu_1(F)$, and thus $\mu_2 \geq \mu_1$ as set functions on $\mathcal{A}$. $\square$

6. Completion of a Measure Space

If $(X, \mathcal{A}, \mu)$ is a measure space, we define the **completion** of this space to be the measure space $(X, \overline{\mathcal{A}}, \overline{\mu})$ defined by

$$\overline{\mathcal{A}} = \left\{ E \,\Delta\, Z \;\middle|\; \begin{array}{l} E \text{ is in } \mathcal{A} \text{ and } Z \subseteq Z' \text{ for} \\ \text{some } Z' \in \mathcal{A} \text{ with } \mu(Z') = 0 \end{array} \right\},$$

$$\overline{\mu}(E \,\Delta\, Z) = \mu(E).$$

It is necessary to verify that the result is in fact a measure space, and we shall carry out this step in the proposition below. In the case of Lebesgue measure m on the line, when initially defined on the σ-algebra $\mathcal{A}$ of Borel sets, the sets in σ-algebra $\overline{\mathcal{A}}$ are said to be **Lebesgue measurable**.

Proposition 5.38. If $(X, \mathcal{A}, \mu)$ is a measure space, then the completion $(X, \overline{\mathcal{A}}, \overline{\mu})$ is a measure space. Specifically

(a) $\overline{\mathcal{A}}$ is a σ-algebra containing $\mathcal{A}$,
(b) the set function $\overline{\mu}$ is unambiguously defined on $\overline{\mathcal{A}}$, i.e., if $E_1 \,\Delta\, Z_1 = E_2 \,\Delta\, Z_2$ as above, then $\mu(E_1) = \mu(E_2)$,
(c) $\overline{\mu}$ is a measure on $\overline{\mathcal{A}}$, and $\overline{\mu}(E) = \mu(E)$ for all sets E in $\mathcal{A}$.

In addition,

(d) if $\widetilde{\mu}$ is any measure on $\overline{\mathcal{A}}$ such that $\widetilde{\mu}(E) = \mu(E)$ for all E in $\mathcal{A}$, then $\widetilde{\mu} = \overline{\mu}$ on $\overline{\mathcal{A}}$,

(e) if $\mu(X) < +\infty$ and if for $E \subseteq X$, $\mu_*(E)$ and $\mu^*(E)$ are defined by

$$\mu_*(E) = \sup_{A \subseteq E,\ A \in \mathcal{A}} \mu(A) \qquad \text{and} \qquad \mu^*(E) = \inf_{A \supseteq E,\ A \in \mathcal{A}} \mu(A),$$

then E is in $\overline{\mathcal{A}}$ if and only if $\mu_*(E) = \mu^*(E)$.

PROOF. For (a), certainly $\mathcal{A} \subseteq \overline{\mathcal{A}}$ because we can use $Z = Z' = \varnothing$ in the definition of $\overline{\mathcal{A}}$. Since $(E \,\Delta\, Z)^c = (E \,\Delta\, Z) \,\Delta\, X = (E \,\Delta\, X) \,\Delta\, Z = E^c \,\Delta\, Z$, $\overline{\mathcal{A}}$ is closed under complements.

To prove closure under countable unions, let us first prove that

$$\overline{\mathcal{A}} = \left\{ E \cup Z \;\middle|\; \begin{array}{l} E \text{ is in } \mathcal{A} \text{ and } Z \subseteq Z' \text{ for} \\ \text{some } Z' \in \mathcal{A} \text{ with } \mu(Z') = 0 \end{array} \right\}. \tag{$*$}$$

Thus let $E \cup Z$ be given, with $Z \subseteq Z'$. Then $E \cup Z = E \,\Delta\, (Z \,\Delta\, (E \cap Z))$ with $Z \,\Delta\, (E \cap Z) \subseteq Z'$. So $E \cup Z$ is in $\overline{\mathcal{A}}$. Conversely if $E \,\Delta\, Z$ is in $\overline{\mathcal{A}}$, we can write $E \,\Delta\, Z = (E - Z') \cup ((E \cap Z') - Z) \cup (Z - E))$ with $((E \cap Z') - Z) \cup (Z - E)) \subseteq Z'$, and then we see that $E \,\Delta\, Z$ is of the form $E'' \cup Z''$ with E'' in $\mathcal{A}$ and $Z'' \subseteq Z'$.

Returning to the proof of closure under countable unions, let $E_n \cup Z_n$ be given in $\overline{\mathcal{A}}$ with $Z_n \subseteq Z_n'$ and $\mu(Z_n') = 0$. Then $\bigcup_n (E_n \cup Z_n) = \big(\bigcup_n E_n\big) \cup \big(\bigcup_n Z_n\big)$ with $\bigcup_n Z_n \subseteq \bigcup_n Z_n'$ and $\mu\big(\bigcup_n Z_n'\big) = 0$. In view of $(*)$, $\overline{\mathcal{A}}$ is therefore closed under countable unions.

For (b), we take as given that $E_1 \,\Delta\, Z_1 = E_2 \,\Delta\, Z_2$ with $Z_1 \subseteq Z_1'$, $Z_2 \subseteq Z_2'$, and $\mu(Z_1') = \mu(Z_2') = 0$. Then $(E_1 \,\Delta\, E_2) \,\Delta\, (Z_1 \,\Delta\, Z_2) = \varnothing$ and hence $E_1 \,\Delta\, E_2 = Z_1 \,\Delta\, Z_2 \subseteq Z_1' \cup Z_2'$. Therefore $\mu(E_1 - E_2) \le \mu(E_1 \,\Delta\, E_2) \le \mu(Z_1' \cup Z_2') = 0$ and similarly $\mu(E_2 - E_1) = 0$. It follows that $\mu(E_1) = \mu(E_1 - E_2) + \mu(E_1 \cap E_2) = \mu(E_1 \cap E_2) = \mu(E_2 - E_1) + \mu(E_1 \cap E_2) = \mu(E_2)$, and $\overline{\mu}$ is unambiguously defined.

For (c), we see from $(*)$ that $\overline{\mu}$ can be defined equivalently by $\overline{\mu}(E \cup Z) = \mu(E)$ if $Z \subseteq Z'$ and $\mu(Z') = 0$. If a disjoint sequence $E_n \cup Z_n$ is given, then we find that $\overline{\mu}\big(\bigcup_n (E_n \cup Z_n)\big) = \overline{\mu}\big(\big(\bigcup_n E_n\big) \cup \big(\bigcup_n Z_n\big)\big) = \mu\big(\bigcup_n E_n\big) = \sum \mu(E_n) = \sum \overline{\mu}(E_n \cup Z_n)$, and complete additivity is proved. Taking $Z = \varnothing$ in the definition $\overline{\mu}(E \cup Z) = \mu(E)$, we obtain $\overline{\mu}(E) = \mu(E)$ for E in $\mathcal{A}$.

For (d), we use $(*)$ as the description of the sets in $\overline{\mathcal{A}}$. Let $E \cup Z$ be in $\overline{\mathcal{A}}$ with E in $\mathcal{A}$, $Z \subseteq Z'$, and Z' in $\mathcal{A}$ with $\mu(Z') = 0$. Then Proposition 5.1e gives $\widetilde{\mu}(E \cap Z) \le \widetilde{\mu}(Z) \le \widetilde{\mu}(Z') = \mu(Z') = 0$, so that $\widetilde{\mu}(E \cap Z) = \widetilde{\mu}(Z) = 0$. Meanwhile, Proposition 5.1b gives $\widetilde{\mu}(E \cup Z) + \widetilde{\mu}(E \cap Z) = \widetilde{\mu}(E) + \widetilde{\mu}(Z)$. Hence $\widetilde{\mu}(E \cup Z) = \widetilde{\mu}(E) = \mu(E)$.

For (e), it is immediate that $\mu_*(E) \le \mu^*(E)$ for every subset E of X. Let $E = C \cup Z$ be in $\overline{\mathcal{A}}$ with C in $\mathcal{A}$, $Z \subseteq Z'$, and Z' in $\mathcal{A}$ with $\mu(Z') = 0$. Then $\mu(C) \le \mu_*(E) \le \mu^*(E) \le \mu(C \cup Z') \le \mu(C) + \mu(Z') = \mu(C)$. Since the expressions at the ends are equal, we must have equality throughout, and therefore $\mu_*(E) = \mu^*(E)$.

In the converse direction let $\mu_*(E) = \mu^*(E)$. We can find a sequence of sets $A_n \in \mathcal{A}$ contained in E with $\lim \mu(A_n) = \mu_*(E)$, and we may assume without loss of generality that $\{A_n\}$ is an increasing sequence. Similarly we can find a decreasing sequence of sets $B_n \in \mathcal{A}$ containing E with $\lim \mu(B_n) = \mu^*(E)$. Let $A = \bigcup_n A_n$ and $B = \bigcap_n B_n$. When combined with the equality $\mu_*(E) = \mu^*(E)$, Proposition 5.2 and Corollary 5.3 show that $\mu(A) = \mu_*(E) = \mu^*(E) = \mu(B)$. Since $A \subseteq E \subseteq B$, we have $\mu(B-A) = \mu(B) - \mu(A) = 0$ and $E = A \cup (E-A)$ with $E - A \subseteq B - A$ and $\mu(B - A) = 0$. By $(*)$, E is in $\overline{\mathcal{A}}$. □

A variant of Proposition 5.38e and its proof identifies the σ-algebra on which the extended measure is constructed in the proof of the Extension Theorem (Theorem 5.5) in the special case we considered. In the special case of the Extension Theorem, the given ring of sets is an algebra $\mathcal{A}$, and $\nu(X)$ is finite. The set function ν gets extended to a measure μ on a σ-algebra $\mathcal{B}$ that contains the smallest σ-algebra $\mathcal{C}$ containing $\mathcal{A}$. The sets of $\mathcal{B}$ are those for which $\mu_*(E) = \mu^*(E)$, where

$$\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U) \qquad \text{and} \qquad \mu_*(E) = \sup_{K \subseteq E,\, K \in \mathcal{K}} \mu_*(K),$$

$\mathcal{K}$ and $\mathcal{U}$ having been defined in terms of countable intersections and countable unions, respectively, from $\mathcal{A}$. The variant of Proposition 5.38e is that a subset E of X has $\mu_*(E) = \mu^*(E)$ if and only if E is of the form $C \cup Z$ with C in $\mathcal{C}$, $Z \subseteq Z'$, and Z' in $\mathcal{C}$ with $\mu(Z') = 0$. In other words, $(X, \mathcal{B}, \mu)$ is the completion of $(X, \mathcal{C}, \mu)$.

The proof is modeled on the proof of Proposition 5.38e. If $E = C \cup Z$ is a set in $\overline{\mathcal{C}}$ with C in $\mathcal{C}$, $Z \subseteq Z'$, and Z' in $\mathcal{C}$ with $\mu(Z') = 0$, then $\mu(C) \le \mu_*(E) \le \mu^*(E) \le \mu(C \cup Z') \le \mu(C) + \mu(Z') = \mu(C)$. We conclude that $\mu_*(E) = \mu^*(E)$.

In the converse direction let $\mu_*(E) = \mu^*(E)$. We can find an increasing sequence of sets $A_n \in \mathcal{K} \subseteq \mathcal{C}$ contained in E with $\lim \mu(A_n) = \mu_*(E)$, and we can find a decreasing sequence of sets $B_n \in \mathcal{U} \subseteq \mathcal{C}$ containing E with $\lim \mu(B_n) = \mu^*(E)$. Let $A = \bigcup_n A_n$ and $B = \bigcap_n B_n$. Arguing as in the proof of Proposition 5.38e, we have $\mu(A) = \mu_*(E) = \mu^*(E) = \mu(B)$, $\mu(B - A) = \mu(B) - \mu(A) = 0$, and $E = A \cup (E - A)$ with $E - A \subseteq B - A$ and $\mu(B - A) = 0$. Thus $E = C \cup Z$ with $C = A$ and $Z = E - A$.

This calculation has the following interesting consequence.

Proposition 5.39. In $\mathbb{R}^1$, the Lebesgue measurable sets of measure 0 are exactly the subsets E of $\mathbb{R}^1$ with the following property: for any $\epsilon > 0$, the set E can be covered by countably many intervals of total length less than ϵ.

PROOF. Within a bounded interval $[a, b]$, the above remarks apply and show that the Lebesgue measurable sets of measure 0 are the sets E with $\mu^*(E) = 0$, where $\mu^*(E) = \inf_{U \supseteq E,\, U \in \mathcal{U}} \mu^*(U)$. The sets U defining $\mu^*(E)$ are countable unions of intervals, and the proposition follows for subsets of any bounded interval $[a, b]$.

For general sets E in $\mathbb{R}^1$, if the covering condition holds, then Proposition 5.1g shows that E has Lebesgue measure 0. Conversely if E is Lebesgue measurable of measure 0, then $E \cap [-N, N]$ is a bounded set of measure 0 and can be covered by countably many intervals of arbitrarily small total length. Let us arrange that the total length is $< 2^{-N}\epsilon$. Taking the union of these sets of intervals as N varies, we obtain a cover of E by countably many intervals of total length less than ϵ. $\square$

7. Fubini's Theorem for the Lebesgue Integral

Fubini's Theorem for the Lebesgue integral concerns the interchange of order of integration of functions of two variables, just as with the Riemann integral in Section III.9. In the case of Euclidean space $\mathbb{R}^n$, we could have constructed Lebesgue measure in each dimension by a procedure similar to the one we used for $\mathbb{R}^1$. Then Fubini's Theorem relates integration of a function of $k + l$ variables over a set by either integrating in all variables at once or integrating in the first k variables first or integrating in the last l variables first. In the context of more general measure spaces, we need to develop the notion of the product of two measure spaces. This corresponds to knowing $\mathbb{R}^k$ and $\mathbb{R}^l$ with their Lebesgue measures and to constructing $\mathbb{R}^{k+l}$ with its Lebesgue measure.

In the theorem as we shall state it, we are given two measures spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, and we assume that both μ and ν are σ-finite. We shall construct a product measure space$(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$, and the formula in question will be

$$\begin{aligned} \int_{X \times Y} f \, d(\mu \times \nu) &\stackrel{?}{=} \int_X \Big[\int_Y f(x, y) \, d\nu(y) \Big] d\mu(x) \\ &\stackrel{?}{=} \int_Y \Big[\int_X f(x, y) \, d\mu(x) \Big] d\nu(y). \end{aligned}$$

This formula will be valid for $f \geq 0$ measurable with respect to $\mathcal{A} \times \mathcal{B}$.

The technique of proof will be the standard one indicated in connection with proving Corollary 5.28. We start with indicator functions, extend the result to simple functions by linearity, and pass to the limit by the Monotone Convergence

Theorem (Theorem 5.25). It is then apparent that the difficult step is the case that f is an indicator function. In fact, it is not even clear in this special case that the inside integral $\int_Y I_E(x, y)\,d\nu(y)$ is a measurable function of X, and this is the step that requires some work.

We begin by describing $\mathcal{A}\times\mathcal{B}$, the σ-algebra of measurable sets for the product $X \times Y$. Recall from Section A1 of Appendix A that $X \times Y$ is defined as a set of ordered pairs. If $A \subseteq X$ and $B \subseteq Y$, then the set of ordered pairs that constitute $A \times B$ is a subset of $X \times Y$, and we call $A \times B$ a **rectangle**[6] in $X \times Y$. The sets A and B are called the **sides** of the rectangle.

Proposition 5.40. If $\mathcal{A}$ and $\mathcal{B}$ are algebras of subsets of nonempty sets X and Y, then the class $\mathcal{C}$ of all finite disjoint unions of rectangles $A \times B$ with A in $\mathcal{A}$ and B in $\mathcal{B}$ is an algebra of sets in $X \times Y$. In particular, a finite union of rectangles is a finite disjoint union.

PROOF. The intersection of the rectangles $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ is the rectangle $R = (A_1 \cap A_2) \times (B_1 \cap B_2)$ because both $R_1 \cap R_2$ and R coincide with the set $\big\{(x, y) \in X \times Y \bigm| x \in A_1,\ x \in A_2,\ y \in B_1,\ y \in B_2\big\}$. Therefore

$$\Big(\bigcup_{i=1}^{m}(A_i \times B_i)\Big) \cap \Big(\bigcup_{j=1}^{n}(C_j \times D_j)\Big) = \bigcup_{i,j} \big\{(A_i \cap C_j) \times (B_i \cap D_j)\big\},$$

and the right side is a disjoint union if both $\bigcup_i (A_i \times B_i)$ and $\bigcup_j (C_j \times D_j)$ are disjoint unions. Moreover, the right side is in $\mathcal{C}$ if both unions on the left are in $\mathcal{C}$. Therefore $\mathcal{C}$ is closed under finite intersections.

Certainly $\varnothing$ and $X \times Y$ are in $\mathcal{C}$. The identity

$$(X \times Y) - (A \times B) = \big((X - A) \times B\big) \cup \big(X \times (Y - B)\big)$$

exhibits the complement of a rectangle as a disjoint union of rectangles. Since the complement of a disjoint union is the intersection of the complements, $\mathcal{C}$ is closed under complementation. Thus $\mathcal{C}$ is an algebra of sets, and the proof is complete. $\square$

If $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras in X and Y, then we denote the smallest σ-algebra containing the algebra $\mathcal{C}$ of the above proposition by $\mathcal{A} \times \mathcal{B}$. The set $X \times Y$, together with the σ-algebra $\mathcal{A} \times \mathcal{B}$, is called a product space. The **measurable sets** of $X \times Y$ are the sets of $\mathcal{A} \times \mathcal{B}$.

[6]The word "rectangle" was used with a different meaning in Chapter III, but there will be no possibility of confusion for now. Starting in Chapter VI, both kinds of rectangles will be in play; the ones in Chapter III can then be called "geometric rectangles" and the present ones can be called "abstract rectangles."

Let E be any set in $X \times Y$. The **section** E_x of E determined by x in X is defined by

$$E_x = \{y \mid (x, y) \text{ is in } E\}.$$

Similarly the section E^y determined by y in Y is

$$E^y = \{x \mid (x, y) \text{ is in } E\}.$$

The section E_x is a subset of Y, and the section E^y is a subset of X.

Lemma 5.41. Let $\{E_\alpha\}$ be a class of subsets of $X \times Y$, and let x be a point of X. Then

(a) $\big(\bigcup_\alpha E_\alpha\big)_x = \bigcup_\alpha (E_\alpha)_x$,
(b) $\big(\bigcap_\alpha E_\alpha\big)_x = \bigcap_\alpha (E_\alpha)_x$,
(c) $(E_\alpha - E_\beta)_x = (E_\alpha)_x - (E_\beta)_x$ and, in particular, $(E_\beta^c)_x = Y - (E_\beta)_x$.

PROOF. These facts are special cases of the identities at the end of Section A1 of Appendix A for inverse images of functions. In this case the function in question is given by $f(y) = (x, y)$. $\square$

Proposition 5.42. Let $\mathcal{A}$ and $\mathcal{B}$ be σ-algebras in X and Y, and let E be a measurable set in $X \times Y$. Then every section E_x is a measurable set in Y, and every section E^y is a measurable set in X.

PROOF. We prove the result for sections E_x, the proof for E^y being completely analogous. Let $\mathcal{E}$ be the class of all subsets E of $X \times Y$ all of whose sections E_x are in $\mathcal{B}$. Then $\mathcal{E}$ contains all rectangles with measurable sides, since a section of a rectangle is either the empty set or one of the sides. By Lemma 5.41a, $\mathcal{E}$ is closed under finite unions. Hence $\mathcal{E}$ contains the algebra $\mathcal{C}$ of finite disjoint unions of rectangles with measurable sides. By parts (a) and (c) of Lemma 5.41, $\mathcal{E}$ is closed under countable unions and complements. It is therefore a σ-algebra containing $\mathcal{C}$ and thus contains $\mathcal{A} \times \mathcal{B}$. $\square$

A corollary of Proposition 5.42 is that a rectangle in $X \times Y$ is measurable if and only if its sides are measurable. The sufficiency follows from the fact that a rectangle with measurable sides is in $\mathcal{C}$, and the necessity follows from the proposition.

From now on, we shall adhere to the convention that a **rectangle** is always assumed to be measurable.

We turn to the implementation of the sketch of proof of Fubini's Theorem given earlier in this section. The basic question will be the equality of the iterated integrals in either order when the integrand is an indicator function. If E is a measurable set in $X \times Y$, then we know from Proposition 5.42 that E_x is a measurable subset of Y. In order to form the iterated integral

$$\int_X \Big[\int_Y I_E(x, y)\, d\nu(y) \Big]\, d\mu(x),$$

we compute the inside integral as $\nu(E_x)$, and we have to be able to form the outside integral, which is $\int_X \nu(E_x)\, d\mu(x)$. That is, we need to know that $\nu(E_x)$ is a measurable function on X. For the iterated integral in the other order, we need to know that $\mu(E^y)$ is measurable on Y.

The proof of this measurability is the hard step, since the class of sets E for which $\nu(E_x)$ and $\mu(E^y)$ are both measurable does not appear to be necessarily a σ-algebra, even when μ and ν are finite measures. To deal with this difficulty, we introduce the following terminology: a class of sets is called a **monotone class** if it is closed under countable increasing unions and countable decreasing intersections. It is readily verified that the class of all subsets of a set is a monotone class and that the intersection of any nonempty family of monotone classes is a monotone class; hence there is a smallest monotone class containing any given class of sets.

The proof of the lemma below introduces the notation $\uparrow$ and $\downarrow$ to denote increasing countable union and decreasing countable intersection, respectively.

Lemma 5.43 (Monotone Class Lemma). The smallest monotone class $\mathcal{M}$ containing an algebra $\mathcal{A}$ of sets is identical to the smallest σ-algebra $\widetilde{\mathcal{A}}$ containing $\mathcal{A}$.

PROOF. We have $\mathcal{M} \subseteq \widetilde{\mathcal{A}}$ because $\widetilde{\mathcal{A}}$ is a monotone class containing $\mathcal{A}$. To prove the reverse inclusion, it is sufficient to show that $\mathcal{M}$ is closed under the operations of finite union and complementation, since a countable union can be written as the increasing countable union of finite unions. The proof is in three steps.

First we prove that if A is in $\mathcal{A}$ and M is in $\mathcal{M}$, then $A \cup M$ and $A \cap M$ are in $\mathcal{M}$. For fixed A in $\mathcal{A}$, let $\mathcal{U}_A$ be the class of all sets M in $\mathcal{M}$ such that $A \cup M$ and $A \cap M$ are in $\mathcal{M}$. Then $\mathcal{U}_A \supseteq \mathcal{A}$. If we show that $\mathcal{U}_A$ is a monotone class, then it will follow that $\mathcal{U}_A \supseteq \mathcal{M}$. For this purpose let

$$U_n \uparrow U \quad \text{and} \quad V_n \downarrow V \quad \text{with} \quad U_n \text{ and } V_n \text{ in } \mathcal{U}_A.$$

By definition of $\mathcal{U}_A$, the sets $U_n \cup A$, $U_n \cap A$, $V_n \cup A$, and $V_n \cap A$ are in $\mathcal{M}$. But

$$U_n \cup A \uparrow U \cup A \quad \text{and} \quad U_n \cap A \uparrow U \cap A,$$
$$V_n \cup A \downarrow V \cup A \quad \text{and} \quad V_n \cap A \downarrow V \cap A.$$

Therefore U and V are in $\mathcal{U}_A$, and $\mathcal{U}_A$ is a monotone class.

Second we prove that $\mathcal{M}$ is closed under finite unions. For fixed N in $\mathcal{M}$, let $\mathcal{U}_N$ be the class of all sets M in $\mathcal{M}$ such that $N \cup M$ and $N \cap M$ are in $\mathcal{M}$. Then $\mathcal{U}_N \supseteq \mathcal{A}$ by the previous step. The same argument as in that step shows that $\mathcal{U}_N$ is a monotone class, and hence $\mathcal{U}_N = \mathcal{M}$.

Third we prove that $\mathcal{M}$ is closed under complements. Let $\mathcal{N}$ be the class of all sets in $\mathcal{M}$ whose complements are in $\mathcal{M}$. Then $\mathcal{N} \supseteq \mathcal{A}$, and it is enough to show that $\mathcal{N}$ is a monotone class. If

$$C_n \uparrow C \quad \text{and} \quad D_n \downarrow D \quad \text{with} \quad C_n \text{ and } D_n \text{ in } \mathcal{N},$$

then C and D are in $\mathcal{M}$ since C_n and D_n are in $\mathcal{M}$. Now

$$C_n^c \downarrow C^c \quad \text{and} \quad D_n^c \uparrow D^c,$$

and by definition of $\mathcal{N}$, C_n^c and D_n^c are in $\mathcal{M}$. Therefore C^c and D^c are in $\mathcal{M}$, and C and D must be in $\mathcal{N}$. That is, $\mathcal{N}$ is a monotone class. □

Lemma 5.44. If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces, then $\nu(E_x)$ and $\mu(E^y)$ are measurable functions for every E in $\mathcal{A} \times \mathcal{B}$.

PROOF IF $\mu(X) < +\infty$ AND $\nu(Y) < +\infty$. Let $\mathcal{M}$ be the class of all sets E in $\mathcal{A} \times \mathcal{B}$ for which $\nu(E_x)$ and $\mu(E^y)$ are measurable. We shall show that $\mathcal{M}$ is a monotone class containing the algebra $\mathcal{C}$ of finite disjoint unions of rectangles. If $R = A \times B$ is a rectangle, then

$$\nu(R_x) = \nu(B) I_A \qquad \text{and} \qquad \mu(R^y) = \mu(A) I_B,$$

and so R is in $\mathcal{M}$. If E and F are disjoint sets in $\mathcal{M}$, then

$$\nu((E \cup F)_x) = \nu(E_x \cup F_x) = \nu(E_x) + \nu(F_x)$$

for each x, and similarly for μ for each y. By Proposition 5.7, $\nu((E \cup F)_x)$ and $\mu((E \cup F)^y)$ are measurable. Hence $E \cup F$ is in $\mathcal{M}$, and $\mathcal{M}$ contains $\mathcal{C}$. If $\{E_n\}$ and $\{F_n\}$ are increasing and decreasing sequences of sets in $\mathcal{M}$, then the finiteness and complete additivity of ν imply that

$$\nu\Big(\Big(\bigcup_n E_n\Big)_x\Big) = \nu\Big(\bigcup_n (E_n)_x\Big) = \lim \nu((E_n)_x)$$

and

$$\nu\Big(\Big(\bigcap_n F_n\Big)_x\Big) = \nu\Big(\bigcap_n (F_n)_x\Big) = \lim \nu((F_n)_x),$$

and similarly for μ. Since the limit of measurable functions is measurable (Corollary 5.10), we conclude that $\mathcal{M}$ is a monotone class. Therefore $\mathcal{M}$ contains $\mathcal{A} \times \mathcal{B}$ by the Monotone Class Lemma (Lemma 5.43). □

PROOF FOR σ-FINITE μ AND ν. Write $X = \bigcup_{m=1}^{\infty} X_m$ and $Y = \bigcup_{n=1}^{\infty} Y_n$ disjointly, with $\mu(X_m) < +\infty$ and $\nu(Y_n) < +\infty$ for all m and n. Define $\mathcal{A}_m$ and $\mathcal{B}_n$ by

$$\mathcal{A}_m = \{A \cap X_m \mid A \text{ is in } \mathcal{A}\} \quad \text{and} \quad \mathcal{B}_n = \{B \cap Y_n \mid B \text{ is in } \mathcal{B}\},$$

and define μ_m and ν_n on $\mathcal{A}_m$ and $\mathcal{B}_n$ by restriction from μ and ν. Then the triples $(X_m, \mathcal{A}_m, \mu_m)$ and $(Y_n, \mathcal{B}_n, \nu_n)$ are finite measure spaces, and the previous case applies. If E is in $\mathcal{A} \times \mathcal{B}$, then $E_{mn} = E \cap (X_m \times Y_n)$ is in $\mathcal{A}_m \times \mathcal{B}_n$, and so $\nu((E_{mn})_x)$ and $\mu((E_{mn})^y)$ are measurable with respect to $\mathcal{A}_m$ and $\mathcal{B}_n$, hence with respect to $\mathcal{A}$ and $\mathcal{B}$. Thus

$$\nu(E_x) = \sum_{m,n} \nu((E_{mn})_x) \qquad \text{and} \qquad \mu(E^y) = \sum_{m,n} \mu((E_{mn})^y)$$

exhibit $\nu(E_x)$ and $\mu(E^y)$ as countable sums of nonnegative measurable functions. They are therefore measurable.

The next proposition simultaneously constructs the product measure and establishes Fubini's Theorem for indicator functions.

Proposition 5.45. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then there exists a unique measure $\mu \times \nu$ on $\mathcal{A} \times \mathcal{B}$ such that

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$$

for every rectangle $A \times B$. The measure $\mu \times \nu$ is σ-finite, and

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y)$$

for every set E in $\mathcal{A} \times \mathcal{B}$.

PROOF. In view of the measurability of $\nu(E_x)$ given in Lemma 5.44, we can define a set function ρ on $\mathcal{A} \times \mathcal{B}$ by

$$\rho(E) = \int_X \nu(E_x)\, d\mu(x).$$

Then $\rho(\varnothing) = 0$, and ρ is nonnegative. On a rectangle $A \times B$, we have

$$\rho(A \times B) = \mu(A)\nu(B) \tag{$*$}$$

since $\nu((A \times B)_x) = \nu(B) I_A$. We shall show that ρ is completely additive. If $\{E_n\}$ is a disjoint sequence in $\mathcal{A} \times \mathcal{B}$, then

$$
\begin{aligned}
\rho\Big(\bigcup_n E_n\Big) &= \int_X \nu\Big(\Big(\bigcup_n E_n\Big)_x\Big)\,d\mu(x) && \text{by definition of } \rho \\
&= \int_X \nu\Big(\bigcup_n (E_n)_x\Big)\,d\mu(x) && \text{by Lemma 5.41a} \\
&= \int_X \Big[\sum_n \nu((E_n)_x)\Big]\,d\mu(x) && \text{since the sets } (E_n)_x \text{ are disjoint for each fixed } x \\
&= \sum_n \int_X \nu((E_n)_x)\,d\mu(x) && \text{by Corollary 5.27} \\
&= \sum_n \rho(E_n).
\end{aligned}
$$

Now $X \times Y = \bigcup_{m,n} (X_m \times Y_n)$. Since ρ has just been shown to be completely additive and since μ and ν are σ-finite, $(*)$ shows that ρ is σ-finite. Also, $(*)$ completely determines ρ on the algebra $\mathcal{C}$ of finite disjoint unions of rectangles. By the Extension Theorem (Theorem 5.5), ρ is completely determined on the smallest σ-algebra $\mathcal{A} \times \mathcal{B}$ containing $\mathcal{C}$.

Defining $\sigma(E) = \int_Y \mu(E^y)\,d\nu(y)$ and arguing in the same way, we see that σ is a measure on $\mathcal{A} \times \mathcal{B}$ agreeing with ρ on rectangles and determined on $\mathcal{A} \times \mathcal{B}$ by its values on rectangles. Thus we have $\rho = \sigma$ on $\mathcal{A} \times \mathcal{B}$, and can define $\mu \times \nu = \rho = \sigma$ to complete the proof. □

Lemma 5.46. If f is a measurable function defined on a product space $X \times Y$, then for each x in X, $y \mapsto f(x, y)$ is a measurable function on Y, and for each y in Y, $x \mapsto f(x, y)$ is a measurable function on X.

PROOF. For each fixed x, the formula

$$\big\{y \,\big|\, f(x, y) > c\big\} = \big\{(x, y) \,\big|\, f(x, y) > c\big\}_x$$

exhibits the set on the left as a section of a measurable set, which must be measurable according to Proposition 5.42. The result for fixed y is proved similarly. □

Theorem 5.47 (Fubini's Theorem). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $(X \times Y, \mathcal{A} \times \mathcal{B}, \mu \times \nu)$ be the product measure space. If f is a *nonnegative measurable* function on $X \times Y$, then $\int_Y f(x, y)\,d\nu(y)$ and $\int_X f(x, y)\,d\mu(x)$ are measurable, and

$$
\begin{aligned}
\int_{X\times Y} f\,d(\mu \times \nu) &= \int_X \Big[\int_Y f(x, y)\,d\nu(y)\Big]\,d\mu(x) \\
&= \int_Y \Big[\int_X f(x, y)\,d\mu(x)\Big]\,d\nu(y).
\end{aligned}
$$

PROOF. Lemma 5.46 shows that $f(x, y)$ is measurable in each variable separately and hence that the inside integrals in the conclusion are well defined. If f is the indicator function of a measurable subset E of $X \times Y$, then the theorem reduces to Proposition 5.45. The result immediately extends to the case of a simple function $f \geq 0$.

Now let f be an arbitrary nonnegative measurable function. Find by Proposition 5.11 an increasing sequence of simple functions $s_n \geq 0$ with pointwise limit f. The sequence of functions $\int_Y s_n(x, y)\, d\nu(y)$ is an increasing sequence of nonnegative functions, and each is measurable by what we have already shown for simple functions. By the Monotone Convergence Theorem (Theorem 5.25),

$$\lim_n \int_Y s_n(x, y)\, d\nu(y) = \int_Y \lim_n s_n(x, y)\, d\nu(y) = \int_Y f(x, y)\, d\nu(y).$$

Therefore $\int_Y f(x, y)\, d\nu(y)$ is the pointwise limit of measurable functions and is measurable. Similarly $\int_X f(x, y)\, d\mu(x)$ is measurable.

For every n, the result for simple functions gives

$$\int_{X \times Y} s_n\, d(\mu \times \nu) = \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big]\, d\mu(x).$$

By a second application of monotone convergence,

$$\int_{X \times Y} f\, d(\mu \times \nu) = \lim_n \int_{X \times Y} s_n\, d(\mu \times \nu) = \lim_n \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big]\, d\mu(x).$$

By a third application of monotone convergence,

$$\lim_n \int_X \Big[\int_Y s_n(x, y)\, d\nu(y) \Big]\, d\mu(x) = \int_X \Big[\lim_n \int_Y s_n(x, y)\, d\nu(y) \Big]\, d\mu(x).$$

Putting our results together, we obtain

$$\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \Big[\int_Y f(x, y)\, d\nu(y) \Big]\, d\mu(x).$$

The other equality of the conclusion follows by interchanging the roles of X and Y. $\square$

Fubini's Theorem arises surprisingly often in practice. In some applications the theorem is applied at least in part to prove that an integral with a parameter is finite or is 0 for almost every value of the parameter. Here is a general result concerning integral 0.

Corollary 5.48. Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ-finite measure spaces, and suppose that E is a measurable subset of $X \times Y$ such that

$$\nu\big(\{y \mid (x, y) \in E\}\big) = 0$$

for almost every x $[d\mu]$. Then $\mu\big(\{x \mid (x, y) \in E\}\big) = 0$ for almost every y $[d\nu]$.

REMARKS. In words, if the x section of E has ν measure 0 for almost every x in X, then the y section of E has μ measure 0 for almost every y in Y. For example, if one-point sets in X and Y have measure 0 and if every x section of E is a finite subset of Y, then for almost every y in Y, the y section of E has measure 0 in X.

PROOF. Apply Fubini's Theorem to I_E. The iterated integrals are equal, and the hypothesis makes one of them be 0. Then the other one must be 0, and the conclusion follows. $\square$

When one tries to drop the hypothesis in Fubini's Theorem that the integrand is nonnegative, some finiteness condition is needed, and the result in the form of Theorem 5.47 is often used to establish this finiteness. Specifically suppose that f is measurable with respect to $\mathcal{A} \times \mathcal{B}$ but is not necessarily nonnegative. The assumption will be that one of the iterated integrals

$$\int_X \Big[\int_Y |f(x, y)|\, d\nu(y)\Big]\, d\mu(x) \quad \text{and} \quad \int_Y \Big[\int_X |f(x, y)|\, d\mu(x)\Big]\, d\nu(y)$$

is finite. Then the conclusions are that

(a) f is integrable with respect to $\mu \times \nu$;
(b) $\int_Y f(x, y)\, d\nu(y)$ is defined for almost every x $[d\mu]$; if it is redefined to be 0 on the exceptional set, then it is measurable and is in fact integrable $[d\mu]$;
(c) a similar conclusion is valid for $\int_X f(x, y)\, d\mu(x)$;
(d) after the redefinitions in (b) and (c), the double integral equals each iterated integral, and the two iterated integrals are equal.

These conclusions follow immediately by applying Fubini's Theorem to f^+ and f^- separately and subtracting. The redefinitions in (b) and (c) are what make the subtractions of integrands everywhere defined.

One final remark is in order: The completion of $\mathcal{A} \times \mathcal{B}$ is not necessarily the same as the product of the completions of $\mathcal{A}$ and $\mathcal{B}$, and thus the statement of Fubini's Theorem requires some modification if completions of measure spaces are to be used. We shall see in the next chapter that Borel sets in Euclidean space behave well under the formation of product spaces, but Lebesgue measurable sets do not. Thus it simplifies matters to stick to integration of Borel-measurable sets in Euclidean space whenever possible.

8. Integration of Complex-Valued and Vector-Valued Functions

Fix a measure space $(X, \mathcal{A}, \mu)$. In this chapter we have worked so far with measurable functions on X whose values are in $\mathbb{R}^*$, dividing them into two classes as far as integration is concerned. One class consists of measurable functions with values in $[0, +\infty]$, and we defined the integral of any such function as a member of $[0, +\infty]$. The other class consists of general measurable functions with values in $\mathbb{R}^*$. The integral in this case can end up being anything in $\mathbb{R}^*$, and there are some such functions for which the integral is not defined.

It is important in the theory to be able to integrate functions whose values are complex numbers or vectors in $\mathbb{R}^m$ or $\mathbb{C}^m$, and it will not be productive to allow the same broad treatment of infinities as was done for general functions with values in $\mathbb{R}^*$. On the other hand, it is desirable to have the flexibility with nonnegative measurable functions of being able to treat infinite values and infinite integrals in the same way as finite values and finite integrals. In order to have two theories, rather than three, once we pass to vector-valued functions, we shall restrict somewhat the theory we have already developed for general functions with values in $\mathbb{R}^*$.

Let us label these two theories of integration as the one for scalar-valued nonnegative measurable functions and the one for *integrable* vector-valued functions. The first of these theories has already been established and needs no change. The second of these theories needs some definitions and comments that in part repeat steps taken with Riemann integration in Sections I.5, III.3, and III.7 and in part are new. In applications of this second theory later, if the term "vector-valued" is not included in a reference to a function either explicitly or by implication, the convention is that the function is scalar-valued.

In the theory for vector-valued functions, we shall be assuming integrability, and the integrability will force the function to have meaningful finite values almost everywhere. Our convention will be that the values are finite *everywhere*. This will not be a serious restriction for any function that can be considered integrable, since we can redefine such a function on a certain set of measure 0 to be 0, and then the condition will be met without any changes in the values of integrals.

Thus let a function $f : X \to \mathbb{C}^m$ be given. Since the function can have its image contained in $\mathbb{R}^m$, we will be handling $\mathbb{R}^m$-valued functions at the same time. Since m can be 1, we will be handling complex-valued functions at the same time. Since the image can be in $\mathbb{R}^m$ and m can be 1, we will at the same time be recasting our theory of real-valued functions whose values are not necessarily nonnegative. We impose the usual Hermitian inner product $(\cdot\,,\,\cdot)$ and norm $|\cdot|$ on $\mathbb{C}^m$.

The function $\bar{f} : X \to \mathbb{C}^m$ is the composition of f followed by complex conjugation in each entry of $\mathbb{C}^m$. We can write $f = \operatorname{Re} f + i \operatorname{Im} f$, where

$\operatorname{Re} f = \frac{1}{2}(f + \bar{f})$ and $\operatorname{Im} f = \frac{1}{2i}(f - \bar{f})$, and then the functions $\operatorname{Re} f$ and $\operatorname{Im} f$ take values in $\mathbb{R}^m$. Following the convention in Section A7 of Appendix A, let $\{u_1, \ldots, u_m\}$ be the standard basis of $\mathbb{R}^m$.

By a basic open set in $\mathbb{C}^m$, we mean a set that is a product in $\mathbb{R}^{2m}$ of bounded open intervals in each coordinate. In symbols, such a set is centered at some $v_0 \in \mathbb{C}^m$, and there are positive numbers ξ_j and η_j such that the set is

$$\left\{v \in \mathbb{C}^m \;\middle|\; |(\operatorname{Re}(v - v_0), u_j)| < \xi_j \text{ and } |(\operatorname{Im}(v - v_0), u_j)| < \eta_j \text{ for } 1 \le j \le m\right\}.$$

We say that $f : X \to \mathbb{C}^m$ is **measurable** if the inverse image under f of each basic open set in $\mathbb{C}^m$ is measurable, i.e., lies in $\mathcal{A}$.

Lemma 5.49. A function $f : X \to \mathbb{C}^m$ is measurable if and only if the inverse image under f of each open set in $\mathbb{C}^m$ is in $\mathcal{A}$.

PROOF. If the stated condition holds, then the inverse image of any *basic* open set is in $\mathcal{A}$, and hence f is measurable. Conversely suppose f is measurable, and let an open set U in $\mathbb{C}^m$ be given. Then U is the union of a sequence of basic open sets U_n, and the measurability of f, in combination with the formula $f^{-1}(U) = \bigcup_n f^{-1}(U_n)$, shows that $f^{-1}(U)$ is in $\mathcal{A}$. $\square$

Proposition 5.50. A function $f : X \to \mathbb{C}^m$ is measurable if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable.

PROOF. In view of Lemma 5.49, we can work with arbitrary open sets in place of basic open sets. If U and V are open sets in $\mathbb{R}^m$, then the product set $U + iV$ is open in $\mathbb{C}^m$, and $f^{-1}(U + iV) = (\operatorname{Re} f)^{-1}(U) \cap (\operatorname{Im} f)^{-1}(V)$. It is immediate that measurability of $\operatorname{Re} f$ and $\operatorname{Im} f$ implies measurability of f. Conversely if we specialize this formula to $V = \mathbb{R}^m$, then we see that measurability of f implies measurability of $\operatorname{Re} f$. Similarly if we specialize to $U = \mathbb{R}^m$, then we see that measurability of f implies measurability of $\operatorname{Im} f$. $\square$

Proposition 5.51. The following conditions on a function $f : X \to \mathbb{C}^m$ are equivalent:

(a) f is measurable,
(b) (f, v) is measurable for each v in $\mathbb{C}^m$,
(c) (f, u_j) is measurable for $1 \le j \le m$.

REMARKS. When infinite-dimensional ranges are used in more advanced texts, (a) is summarized by saying that f is "strongly measurable," and (b) is summarized by saying that f is "weakly measurable."

PROOF. Suppose (a) holds. The function in (b) is the composition of f followed by the continuous function $(\,\cdot\,, v)$ from $\mathbb{C}^m$ to $\mathbb{C}$. The inverse image of an open set in $\mathbb{C}$ is then open in $\mathbb{C}^m$, and the inverse image of the latter open set under f is in $\mathcal{A}$. This proves (b). Condition (b) trivially implies condition (c). If (c) holds, then Proposition 5.50 shows for each standard basis vector u_j that $(\operatorname{Re} f, u_j)$ and $(\operatorname{Im} f, u_j)$ are measurable from X into $\mathbb{R}$. Thus the inverse image of any open interval under any of these $2m$ functions on X is in $\mathcal{A}$. The inverse image of a basic open set in $\mathbb{C}^m$ under f is the intersection of $2m$ such sets in $\mathcal{A}$ and is therefore in $\mathcal{A}$. Hence (a) holds. □

Proposition 5.52. Measurability of vector-valued functions has the following properties:

(a) If $f : X \to \mathbb{C}^m$ and $g : X \to \mathbb{C}^m$ are measurable, then so is $f + g$ as a function from X to $\mathbb{C}^m$.

(b) If $f : X \to \mathbb{C}^m$ is measurable and c is in $\mathbb{C}$, then cf is measurable as a function from X to $\mathbb{C}^m$.

(c) If $f : X \to \mathbb{C}^m$ is measurable, then so is $\bar{f} : X \to \mathbb{C}^m$.

(d) If $f : X \to \mathbb{C}$ and $g : X \to \mathbb{C}$ are measurable, then so is $fg : X \to \mathbb{C}$.

(e) If $f : X \to \mathbb{C}^m$ is measurable, then $|f| : X \to [0, +\infty)$ is measurable.

(f) If $\{f_n\}$ is a sequence of measurable functions from X into $\mathbb{C}^m$ converging pointwise to a function $f : X \to \mathbb{C}^m$, then f is measurable.

PROOF. Conclusions (a) through (e) may all be proved in the same way. It will be enough to illustrate the technique with (a). We can write the function $x \mapsto f(x) + g(x)$ as a composition of $x \mapsto (f(x), g(x))$ followed by addition $(a, b) \mapsto a+b$. Let an open set in $\mathbb{C}^m$ be given. The inverse image under addition is open in $\mathbb{C}^m \times \mathbb{C}^m$, since addition is continuous (Proposition 2.28). The inverse image of a product $U \times V$ of open sets in $\mathbb{C}^m \times \mathbb{C}^m$ under $x \mapsto (f(x), g(x))$ is $f^{-1}(U) \cap g^{-1}(V)$, which is in $\mathcal{A}$ because f and g are measurable, and therefore the inverse image of any open set in $\mathbb{C}^m \times \mathbb{C}^m$ under $x \mapsto (f(x), g(x))$ is in $\mathcal{A}$. This handles (a), and (b) through (e) are similar.

For (f), we apply Proposition 5.50 to f, and then we apply the equivalence of (a) and (c) of Proposition 5.51 for $\operatorname{Re} f$ and $\operatorname{Im} f$. In this way the result is reduced to the real-valued scalar case, which is known from Corollary 5.10. □

If E is a measurable subset of X, we say that a function $f : X \to \mathbb{C}$ is **integrable** on E if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable on E, and in this case we define $\int_E f\,d\mu = \int_E \operatorname{Re} f\,d\mu + i \int_E \operatorname{Im} f\,d\mu$.

Proposition 5.53. Let E be a measurable subset of X. Integrability on E of functions from X to $\mathbb{C}$ has the following properties:

(a) If f and g are functions from X into $\mathbb{C}$ that are integrable on E, then $f+g$ is integrable on E, and $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.

(b) If f is a function from X into $\mathbb{C}$ that is integrable on E and if c is in $\mathbb{C}$, then cf is integrable on E, and $\int_E cf\,d\mu = c\int_E f\,d\mu$.

(c) If f is a measurable function from X into $\mathbb{C}$ such that $|f|$ is integrable on E, then f is integrable on E, and $\big|\int_E f(x)\,d\mu(x)\big| \le \int_E |f(x)|\,d\mu(x)$.

(d) (Dominated convergence) Let f_n be a sequence of measurable functions from X into $\mathbb{C}$ integrable on E and converging pointwise to f. If there is a measurable function $g : X \to [0,+\infty]$ that is integrable on E and has $|f_n(x)| \le g(x)$ for all x in E, then f is integrable on E, $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{C}$, and $\lim_n \int_E f_n\,d\mu = \int_E f\,d\mu$.

PROOF. Conclusion (a) is immediate from the definitions, and so is (b) for real scalars. Taking (a) and (b) into account, we see that (b) holds if it holds for $c = i$. We have $if = -\operatorname{Im} f + i\operatorname{Re} f$. If f is integrable, then $-\operatorname{Im} f$ and $\operatorname{Re} f$ are integrable, and hence if is integrable. Then

$$\begin{aligned} i\int_E f\,d\mu &= i\big(\int_E \operatorname{Re} f\,d\mu + i\int_E \operatorname{Im} f\,d\mu\big)\\ &= \int_E (-\operatorname{Im} f)\,d\mu + \int_E (i\operatorname{Re} f)\,d\mu = \int_E if\,d\mu, \end{aligned}$$

and hence (b) is proved.

In (c), if $f : X \to \mathbb{C}$ is integrable, choose c with $|c| = 1$ such that $c\int_E f\,d\mu$ is real and ≥ 0. Application of (b) and Proposition 5.16 gives $\big|\int_E f\,d\mu\big| = c\int_E f\,d\mu = \int_E cf\,d\mu = \int_E \operatorname{Re}(cf)\,d\mu \le \int_E |cf|\,d\mu = \int_E |f|\,d\mu$.

Finally (d) follows by applying the Dominated Convergence Theorem (Theorem 5.30) to $\operatorname{Re} f_n$ and $\operatorname{Im} f_n$ separately and then combining the results. □

We turn now to the matter of integrability of vector-valued functions, together with the value of the integral. One way of proceeding is to go back and adapt the theory in Sections 3–4 to work directly with vector-valued functions and approximations by vector-valued simple functions. This approach is useful if at some stage one wants systematically to allow infinite-dimensional vectors as values. Examples of this situation will arise in this book, but there are not enough examples to justify an abstract treatment. One important example arises in the next section with functions of the form $f(x,y)$, which can be regarded as functions of x that take values in a space of functions of y.

Thus we use an abstract definition of integrability that is appropriate only to the case of finite-dimensional range. If E is a measurable subset of X, we say that a function $f : X \to \mathbb{C}^m$ is **integrable** on E if the complex-valued functions (f, u_j) are integrable on E for each u_j in the standard basis, and in this case we define $\int_E f\,d\mu = \sum_{j=1}^m \big(\int_E (f,u_j)\,d\mu\big)u_j$.

Proposition 5.54. Let E be a measurable subset of X. Integrability of vector-valued functions on E satisfies the following properties:

(a) If f and g are functions from X into $\mathbb{C}^m$ that are integrable on E, then $f+g$ is integrable on E, and $\int_X (f+g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu$.

(b) If f is a function from X into $\mathbb{C}^m$ that is integrable on E, then cf is integrable on E, and $\int_E cf\,d\mu = c\int_E f\,d\mu$.

(c) A function $f : X \to \mathbb{C}^m$ is integrable on E if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are integrable on E, and then $\int_X f\,d\mu = \int_X \operatorname{Re} f\,d\mu + i\int_X \operatorname{Im} f\,d\mu$.

(d) If f is a function from X into $\mathbb{C}^m$ that is integrable on E and if v is a member of $\mathbb{C}^m$, then $x \mapsto (f(x), v)$ is integrable on E and $\int_E (f(x), v)\,d\mu(x) = \big(\int_E f(x)\,d\mu(x), v\big)$.

(e) If f is a measurable function from X into $\mathbb{C}^m$ such that $|f|$ is integrable on E, then f is integrable on E, and $\big|\int_E f(x)\,d\mu(x)\big| \le \int_E |f(x)|\,d\mu(x)$.

(f) (Dominated convergence) Let f_n be a sequence of measurable functions from X into $\mathbb{C}^m$ integrable on E and converging pointwise to f. If there is a measurable function $g : X \to [0,+\infty]$ that is integrable on E and has $|f_n(x)| \le g(x)$ for all x in E, then f is integrable on E, $\lim_n \int_E f_n\,d\mu$ exists in $\mathbb{C}^m$, and $\lim_n \int_E f_n\,d\mu = \int_E f\,d\mu$.

PROOF. All of the relevant questions about measurability are addressed by Propositions 5.50 and 5.52. Conclusions (a), (b), (c), and (f) about integrability are immediate from Proposition 5.53.

For (d), let $v = \sum c_j u_j$ with each c_j in $\mathbb{C}$. Since f is by assumption integrable, $(f, v) = (f, \sum c_j u_j) = \sum_j \bar c_j (f, u_j)$ exhibits (f, v) as a linear combination of functions integrable on E. Therefore (f, v) is integrable on E. To obtain the formula asserted in (d), we first consider $v = u_i$. Then the definition of $\int_E f\,d\mu$ gives $\big(\int_E f\,d\mu, u_i\big) = \big(\sum_j \big(\int_E (f, u_j)\,d\mu\big)u_j,\, u_i\big) = \int_E (f, u_i)\,d\mu$. Multiplying by $\bar c_i$ and adding, we obtain $\big(\int_E f\,d\mu, v\big) = \int_E (f, v)\,d\mu$. This proves (d).

For (e), let $f : X \to \mathbb{C}^m$ be measurable on X with $|f|$ integrable on E. The asserted inequality is trivial if $\int_E f\,d\mu = 0$. Otherwise, for every v in $\mathbb{C}^m$,

$$\begin{aligned}
\big|\big(\textstyle\int_E f\,d\mu, v\big)\big| &= \big|\textstyle\int_E (f, v)\,d\mu\big| && \text{by (d)}\\
&\le \textstyle\int_E |(f, v)|\,d\mu && \text{by Proposition 5.53c}\\
&\le |v|\textstyle\int_E |f|\,d\mu && \text{by Proposition 5.16 and the Schwarz inequality.}
\end{aligned}$$

Taking $v = \int_E f\,d\mu$ gives $\big|\int_E f\,d\mu\big|^2 \le \big|\int_E f\,d\mu\big| \int_E |f|\,d\mu$. Since $\int_E f\,d\mu$ has been assumed nonzero, we can divide by its magnitude, and then (e) follows. $\square$

9. L^1, L^2, L^∞, and Normed Linear Spaces

Let $(X, \mathcal{A}, \mu)$ be a measure space. In this section we introduce the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$. Roughly speaking, these will be vector spaces of functions on X with suitable integrability properties. More precisely the actual vector spaces of functions will form pseudometric spaces, and the spaces $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ will be the corresponding metric spaces obtained from the construction of Proposition 2.12. They will all turn out to be vector spaces over $\mathbb{R}$ or $\mathbb{C}$. It will matter little whether the scalars for these vector spaces are real or complex. When we need to refer to operations with scalars, we may use the symbol $\mathbb{F}$ to denote $\mathbb{R}$ or $\mathbb{C}$, and we call $\mathbb{F}$ the **field of scalars**. We shall make explicit mention of $\mathbb{R}$ or $\mathbb{C}$ in any situation in which it is necessary to insist on a particular one of $\mathbb{R}$ or $\mathbb{C}$.

The three spaces we will construct will all be obtained by introducing "pseudonorms" in vector spaces of measurable functions. A **pseudonorm** on a vector space V is a function $\|\cdot\|$ from V to $[0, +\infty)$ such that[7]

(i) $\|x\| \geq 0$ for all $x \in V$,

(ii) $\|cx\| = |c|\|x\|$ for all scalars c and all $x \in V$,

(iii) **(triangle inequality)** $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in V.

We encountered pseudonorms earlier in connection with pseudo inner-product spaces; in Proposition 2.3 we saw how to form a pseudonorm from a pseudo inner product. However, only the pseudonorm for $L^2(X)$ arises from a pseudo inner product in the construction of L^1, L^2, and L^∞.

The definitions of the pseudonorms in these three instances are

$$\begin{aligned} \|f\|_1 &= \textstyle\int_X |f|\,d\mu && \text{for } L^1(X),\\ \|f\|_2 &= \big(\textstyle\int_X |f|^2\,d\mu\big)^{1/2} && \text{for } L^2(X),\\ \|f\|_\infty &= \text{“essential supremum” of } f && \text{for } L^\infty(X). \end{aligned}$$

Once we have defined "essential supremum," all the above expressions are meaningful for any measurable function f from X to the scalars, and the vector space V in each of the cases is the space of all measurable functions from X to the scalars such that the indicated pseudonorm is finite. In other words, V consists of the integrable functions on X in the case of $L^1(X)$, the square-integrable functions on X in the case of $L^2(X)$, and the "essentially bounded" functions on X in the case of $L^\infty(X)$.

We need to check that $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are indeed pseudonorms and that the spaces V are vector spaces in each case.

[7]The word "seminorm" is a second name for a function with these properties and is generally used in the context of a family of such functions. We shall not use the word "seminorm" in this text.

For $L^1(X)$, properties (i) and (ii) are immediate from the definition. For (iii), we have $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ for all x and therefore $\|f + g\|_1 = \int_X |f + g|\,d\mu \leq \int_X |f|\,d\mu + \int_X |g|\,d\mu = \|f\|_1 + \|g\|_1$.

For $L^2(X)$, let V be the space of all square-integrable functions on X. The space V is certainly closed under scalar multiplication; let us see that it is closed under addition. If f and g are in V, then we have

$$\begin{aligned}(|f(x)| + |g(x)|)^2 &\leq \big(\max\{|f(x)|, |g(x)|\} + \max\{|f(x)|, |g(x)|\}\big)^2 \\ &= 4\max\{|f(x)|^2, |g(x)|^2\} \leq 4|f(x)|^2 + 4|g(x)|^2\end{aligned}$$

for every x in X. Integrating over X, we see that $f + g$ is in V if f and g are in V. Also, the left side is $\geq 4|f(x)|\,|g(x)|$, and it follows that $f\bar{g}$ is integrable whenever f and g are in V. Then the definition $(f, g)_2 = \int_E f\bar{g}\,d\mu$ makes V into a pseudo inner product-space in the sense of Section II.1. Hence Proposition 2.3 shows that the function $\|\cdot\|_2$ with $\|f\|_2 = (f, f)_2^{1/2}$ is a pseudonorm on V.

For $L^\infty(X)$, we say that f is **essentially bounded** if there is a real number M such that $|f(x)| \leq M$ almost everywhere $[d\mu]$. Let us call such an M an **essential bound** for $|f|$. When f is essentially bounded, we define $\|f\|_\infty$ to be the infimum of all essential bounds for $|f|$. This infimum is itself an essential bound, since the countable union of sets of measure 0 is of measure 0. The infimum of the essential bounds is called the **essential supremum** of $|f|$. Certainly $\|\cdot\|_\infty$ satisfies (i) and (ii). If $|f|$ is bounded a.e. by M and if $|g|$ is bounded a.e. by N, then $|f + g|$ is bounded everywhere by $|f| + |g|$, which is bounded a.e. by $M + N$. It follows that $f + g$ is essentially bounded and $\|f + g\|_\infty \leq \||f| + |g|\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. So (iii) holds for $\|\cdot\|_\infty$.

A real or complex vector space with a pseudonorm is a **pseudo normed linear space**. Such a space V becomes a pseudometric space by the definition $d(f, g) = \|f - g\|$, according to the proof of Proposition 2.3. Proposition 2.12 shows that if we define two members f and g of V to be equivalent whenever $d(f, g) = 0$, then the result is an equivalence relation and the function d descends to a well-defined metric on the set of equivalence classes. If we take into account the vector space structure on V, then we can see that the operations of addition and scalar multiplication descend to the set of equivalence classes, and the set of equivalence classes is then also a vector space. The argument for addition is that if $d(f_1, f_2) = 0$ and $d(g_1, g_2) = 0$, then $d(f_1 + g_1, f_2 + g_2)$ is 0 because

$$\begin{aligned}d(f_1 + g_1, f_2 + g_2) &= \|(f_1 + g_1) - (f_2 + g_2)\| = \|(f_1 - f_2) + (g_1 - g_2)\| \\ &\leq \|f_1 - f_2\| + \|g_1 - g_2\| = d(f_1, f_2) + d(g_1, g_2) = 0.\end{aligned}$$

The argument for scalar multiplication is similar, and one readily checks that the space of equivalence classes is a vector space.

This construction is to be applied to the spaces V we formed in connection with integrability, square integrability, and essential boundedness. The spaces of equivalence classes in the respective cases are called $L^1(X)$, $L^2(X)$, and $L^\infty(X)$. These spaces of equivalence classes are pseudo normed linear spaces with the additional property that $\|f\| = 0$ only for the 0 element of the vector space. If there is any possibility of confusion, we may write $L^1(\mu)$ or $L^1(X, \mu)$ or $L^1(X, \mathcal{A}, \mu)$ in place of $L^1(X)$, and similarly for L^2 and L^∞.

A pseudo normed linear space is called a **normed linear space** if $\|f\| = 0$ implies f is the 0 element of the vector space. Thus $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ are normed linear spaces.

In practice, in order to avoid clumsiness, one sometimes relaxes the terminology and works with the members of $L^1(X)$, $L^2(X)$, and $L^\infty(X)$ as if they were functions, saying, "Let the function f be in $L^1(X)$" or "Let f be an L^1 function." There is little possibility of ambiguity in using such expressions.

The 1-dimensional vector space consisting of the field of scalars $\mathbb{F}$ with absolute value as norm is an example of a normed linear space. Apart from this and $\mathbb{F}^m$, we have encountered one other important normed linear space thus far in the book. This is the space $B(S)$ of bounded functions on a nonempty set S. It has various vector subspaces of interest, such as the space $C(S)$ of bounded continuous functions in the case that S is a metric space. The norm for $B(S)$ is the **supremum norm** or the **uniform norm** defined by

$$\|f\|_{\sup} = \sup_{s \in S} |f(s)|.$$

The corresponding metric is

$$d(f, g) = \|f - g\|_{\sup} = \sup_{s \in S} |f(s) - g(s)|,$$

and this agrees with the definition of the metric in the example in Chapter II. Proposition 2.44 shows that the metric space $B(S)$ is complete. Any vector subspace of $B(S)$ is a normed linear space under the restriction of the supremum norm to the subspace.

In working with specific normed linear spaces, we shall often be interested in seeing whether a particular subset of the space is dense. In checking denseness, the following proposition about an arbitrary normed linear space is sometimes helpful. The intersection of vector subspaces of X is a vector subspace, and the intersection of closed sets is closed. Therefore it makes sense to speak of the **smallest closed vector subspace** containing a given subset S of X.

Proposition 5.55. If X is a normed linear space with norm $\|\cdot\|$ and with $\mathbb{F}$ as field of scalars, then

(a) addition is a continuous function from $X \times X$ to X,
(b) scalar multiplication is a continuous function from $\mathbb{F} \times X$ to X,

(c) the closure of any vector subspace of X is a vector subspace,
(d) the set of all finite linear combinations of members of a subset S of X is dense in the smallest closed vector subspace containing S.

PROOF. The formula $\|(x+y)-(x_0+y_0)\| \le \|x-x_0\|+\|y-y_0\|$ shows continuity of addition because it says that if x is within distance $\epsilon/2$ of x_0 and y is within distance $\epsilon/2$ of y_0, then $x+y$ is within distance ϵ of x_0+y_0. Similarly the formula $\|cx-c_0x_0\| \le \|c(x-x_0)\|+\|(c-c_0)x_0\| = |c|\|x-x_0\|+|c-c_0|\|x_0\|$ shows that $\|cx-c_0x_0\| \le \delta(|c_0|+1)+\delta\|x_0\|$ as soon as $\delta \le 1$, $|c-c_0| \le \delta$, and $\|x-x_0\| \le \delta$. If ϵ with $0<\epsilon\le 1$ is given and if we set $\delta = (|c_0|+1+\|x_0\|)^{-1}\epsilon$, then we see that $|c-c_0|\le\delta$ and $\|x-x_0\|\le\delta$ together imply $\|cx-c_0x_0\|\le\epsilon$. Hence scalar multiplication is continuous. This proves (a) and (b).

From (a) and (b) it follows that if $x_n \to x$ and $y_n \to y$ in X and $c_n \to c$ in $\mathbb{F}$, then $x_n+y_n \to x+y$ and $c_nx_n \to cx$. This proves (c).

For (d), the smallest closed vector subspace V_1 containing S certainly contains the closure V_2 of the set of all finite linear combinations of members of S. Part (c) shows that V_2 is a closed vector subspace, and hence the definition of V_1 implies that V_1 is contained in V_2. Therefore $V_1 = V_2$, and (d) is proved. □

Proposition 5.56. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p=1$ or $p=2$. Then every indicator function of a set of finite measure is in $L^p(X)$, and the smallest closed subspace of $L^p(X)$ containing all such indicator functions is $L^p(X)$ itself.

REMARK. Proposition 5.55d allows us to conclude from this that the the set of simple functions built from sets of finite measure lies in both $L^1(X)$ and $L^2(X)$ and is dense in each. It of course lies in $L^\infty(X)$ as well, but it is dense in $L^\infty(X)$ if and only if $\mu(X)$ is finite.

PROOF. If E is a set of finite measure, then the equality $\int_X (I_E)^p\,d\mu = \mu(E)$ shows that I_E is in L^p for $p=1$ and $p=2$.

In the reverse direction let V be the smallest closed vector subspace of L^p containing all indicator functions of sets of finite measure. Suppose that $s = \sum_k c_k I_{E_k}$ is the canonical expansion of a simple function $s \ge 0$ in L^p and that $c_k > 0$. The inequalities $0 \le c_k I_{E_k} \le s$ imply that $c_k I_{E_k}$ is in L^p. Hence I_{E_k} is in L^p, and $\mu(E_k)$ is finite. Thus every nonnegative simple function in L^p lies in V.

Let $f \ge 0$ be in L^p, and let s_n be an increasing sequence of simple functions ≥ 0 with pointwise limit f. Since $0 \le s_n \le f$, each s_n is in L^p. Since $|f-s_n|^p$ has pointwise limit 0 and is dominated pointwise for every n by the integrable function $|f|^p$, dominated convergence gives $\lim \int_X |f-s_n|^p\,d\mu = 0$. Hence s_n tends to f in L^p. Combining this conclusion with the result of the previous paragraph, we see that every nonnegative L^p function is in V. Any L^p function is a finite linear combination of nonnegative L^p functions, and hence every L^p function lies in V. $\square$

Let us digress briefly once more from our study of L^1, L^2, and L^∞ to obtain two more results about general normed linear spaces. A linear function between two normed linear spaces is often called a **linear operator**. A linear function whose range space is the field of scalars is called a **linear functional**. The following equivalence of properties is fundamental and is often used without specific reference.

Proposition 5.57. Let X and Y be normed linear spaces that are both real or both complex, and let their respective norms be $\|\cdot\|_X$ and $\|\cdot\|_Y$. Then the following conditions on a linear operator $L : X \to Y$ are equivalent:

(a) L is uniformly continuous on X,
(b) L is continuous on X,
(c) L is continuous at 0,
(d) L is **bounded** in the sense that there exists a constant M such that

$$\|L(x)\|_Y \leq M\|x\|_X$$

for all x in X.

PROOF. If L is uniformly continuous on X, then L is certainly continuous on X. If L is continuous on X, then L is certainly continuous at 0. Thus (a) implies (b), and (b) implies (c).

If L is continuous at 0, find $\delta > 0$ for $\epsilon = 1$ such that $\|x - 0\|_X \leq \delta$ implies $\|L(x) - L(0)\|_Y \leq 1$. Here $L(0) = 0$. If a general $x \neq 0$ is given, then $\|x\|_X \neq 0$, and the properties of the norm give $\|(\delta/\|x\|_X)x\|_X = \delta$. Thus $\|L((\delta/\|x\|_X)x)\|_Y \leq 1$. By the linearity of L and the properties of the norm, $(\delta/\|x\|_X)\|L(x)\|_Y \leq 1$. Therefore $\|L(x)\|_Y \leq \delta^{-1}\|x\|_X$, and L is bounded with $M = \delta^{-1}$. Thus (c) implies (d).

If L is bounded with constant M and if $\epsilon > 0$ is given, let $\delta = \epsilon/M$. Then $\|x_1 - x_2\|_X \leq \delta$ implies

$$\|L(x_1) - L(x_2)\|_Y = \|L(x_1 - x_2)\|_Y \leq M\|x_1 - x_2\|_X \leq \delta M = \epsilon.$$

Thus (d) implies (a). $\square$

If $L : X \to Y$ is a bounded linear operator, then the infimum of all constants M such that $\|L(x)\|_Y \leq M\|x\|_X$ for all x in X is again such a constant, and it is called the **operator norm** $\|L\|$ of L. Thus it in particular satisfies

$$\|L(x)\|_Y \leq \|L\|\|x\|_X \qquad \text{for all } x \text{ in } X.$$

As a consequence of the way that L and the norms in X and Y interact with scalar multiplication, the operator norm is given by the formulas

$$\|L\| = \sup_{\|x\|_X \le 1} \|L(x)\|_Y = \sup_{\|x\|_X = 1} \|L(x)\|_Y$$

except in the uninteresting case $X = 0$. It is easy to check that the bounded linear operators from X into Y form a vector space, and the operator norm makes this vector space into a normed linear space that we denote by $\mathcal{B}(X, Y)$. When the domain and range are the same space X, we refer to the members of $\mathcal{B}(X, X)$ as bounded linear operators *on* X. The normed linear space $\mathcal{B}(X, X)$ has a multiplication operation given by composition.

When Y is the field of scalars $\mathbb{F}$, the space $\mathcal{B}(X, \mathbb{F})$ reduces to the space of continuous linear functionals on X. This is called the **dual space** of X and is denoted by X^*. For example, if $X = L^1(\mu)$, then every member g of $L^\infty(\mu)$ defines a member x_g^* of X^* by $x_g^*(f) = \int fg\,d\mu$ for f in $L^1(\mu)$; the linear functional x_g^* has $\|x_g^*\| \le \|g\|_\infty$. We shall be interested in two kinds of convergence in X^*. One is **norm convergence**, in which a sequence $\{x_n^*\}$ converges to an element x^* in X^* if $\|x_n^* - x^*\|$ tends to 0. The other is **weak-star convergence**, in which $\{x_n^*\}$ converges to x^* weak-star against X if $\lim_n x_n^*(x) = x^*(x)$ for each x in X.

Theorem 5.58 (Alaoglu's Theorem, preliminary form). If X is a separable normed linear space, then any sequence in X^* that is *bounded in norm* has a subsequence that converges weak-star against X.

REMARKS. In Chapter VI we shall see that L^1 and L^2 are separable in the case of Lebesgue measure on $\mathbb{R}^1$ and in the case of many generalizations of Lebesgue measure to N-dimensional Euclidean space.

PROOF. Let a sequence $\{x_n^*\}_{n=1}^\infty$ be given with $\|x_n^*\| \le M$, and let $\{x_k\}$ be a countable dense set in X. For each k, we have $|x_n^*(x_k)| \le \|x_n^*\|\|x_k\| \le M\|x_k\|$, and hence the sequence $\{x_n^*(x_k)\}_{n=1}^\infty$ of scalars is bounded for each fixed k. By the Bolzano–Weierstrass Theorem, $\{x_n^*(x_k)\}_{n=1}^\infty$ has a convergent subsequence. Since we can pass to a convergent subsequence of any subsequence for any particular k, we can use a diagonal process to pass to a single convergent subsequence $\{x_{n_l}^*\}_{l=1}^\infty$ such that $\lim_l x_{n_l}^*(x_k)$ exists for all k.

Now let x_0 be arbitrary in X, let $\epsilon > 0$ be given, and choose x_k in the dense set with $\|x_k - x_0\| < \epsilon$. Then

$$\begin{aligned} |x_{n_l}^*(x_0) - x_{n_{l'}}^*(x_0)| &\le |x_{n_l}^*(x_0 - x_k)| + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)| + |x_{n_{l'}}^*(x_k - x_0)| \\ &\le M\|x_0 - x_k\| + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)| + M\|x_k - x_0\| \\ &\le 2M\epsilon + |x_{n_l}^*(x_k) - x_{n_{l'}}^*(x_k)|. \end{aligned}$$

Thus $\limsup_{l,l'\to\infty} |x^*_{n_l}(x_0) - x^*_{n_{l'}}(x_0)| \le 2M\epsilon$. Since ϵ is arbitrary, we conclude that $\{x^*_{n_l}(x_0)\}_{l=1}^\infty$ is a Cauchy sequence of scalars. It is therefore convergent. Denote the limit by $x^*(x_0)$, so that $\lim_l x^*_{n_l}(x_0) = x^*(x_0)$ for all x_0 in X. Since limits respect addition and multiplication of scalars, x^* is a linear functional on X. The computation $|x^*(x_0)| = |\lim_l x^*_{n_l}(x_0)| = \lim_l |x^*_{n_l}(x_0)| \le \limsup_l \|x^*_l\|\|x_0\| \le M\|x_0\|$ shows that x^* is bounded. Hence $\{x^*_{n_l}\}_{l=1}^\infty$ converges to x^* weak-star against X. □

Now, as promised, we return to L^1, L^2, and L^∞. The completeness asserted in the next theorem will turn out to be one of the key advantages of Lebesgue integration over Riemann integration.

Theorem 5.59. Let $(X, \mathcal{A}, \mu)$ be any measure space, and let p be 1, 2, or ∞. Any Cauchy sequence $\{f_k\}$ in L^p has a subsequence $\{f_{k_n}\}$ such that $\|f_{k_n} - f_{k_m}\|_p \le C_{\min\{m,n\}}$ with $\sum_n C_n < +\infty$. A subsequence $\{f_{k_n}\}$ with this property is necessarily Cauchy pointwise almost everywhere. If f denotes the almost-everywhere limit of $\{f_{n_k}\}$, then the original sequence $\{f_k\}$ converges to f in L^p. Consequently these three spaces L^p, when regarded as metric spaces, are complete in the sense that every Cauchy sequence converges.

REMARKS. The broad sweep of the theorem is that the spaces L^1, L^2, and L^∞ are complete. But the detail is important, too. First of all, the detail allows us to conclude that a sequence convergent in one of these spaces has a subsequence that converges pointwise almost everywhere. Second of all, the detail allows us to conclude that if a sequence of functions is convergent in L^{p_1} and in L^{p_2}, then the limit functions in the two spaces are equal almost everywhere.

PROOF. Let $\{f_n\}$ be a Cauchy sequence in L^p. Inductively choose integers n_k by defining $n_0 = 1$ and taking n_k to be any integer $> n_{k-1}$ such that $\|f_m - f_{n_k}\|_p \le 2^{-k}$ for $m \ge n_k$; we can do so since the given sequence is Cauchy. Then the subsequence $\{f_{n_k}\}$ has the property that $\|f_{n_k} - f_{n_l}\| \le 2^{-\min\{k,l\}}$ for all $k \ge 1$ and $l \ge 1$. This proves the first conclusion of the theorem.

Now suppose that we have a sequence $\{f_n\}$ in L^p such that $\|f_n - f_m\|_p \le C_{\min\{m,n\}}$ with $\sum_n C_n = C < +\infty$. We shall prove that $\{f_n\}$ is Cauchy pointwise almost everywhere and that if f is its almost-everywhere limit, then f_n tends to f in L^p.

First suppose that $p < \infty$. Let g_n be the function from X to $[0, +\infty]$ given by

$$g_n = |f_1| + \sum_{k=2}^{n} |f_k - f_{k-1}|, \qquad (*)$$

and define $g(x) = \lim g_n(x)$ pointwise. Then

$$\begin{aligned}\left(\int_X g_n^p\,d\mu\right)^{1/p} = \|g_n\|_p &\le \|f_1\|_p + \sum_{k=2}^{n} \|f_k - f_{k-1}\|_p \\ &\le \|f_1\|_p + \sum_{k=2}^{n} C_{k-1} \le \|f_1\|_p + C.\end{aligned}$$

By monotone convergence, we deduce that $\left(\int_X g^p\,d\mu\right)^{1/p} = \|g\|_p$ is finite. Thus g is finite a.e., and consequently the series

$$\sum_{k=2}^{\infty} |f_k(x) - f_{k-1}(x)| \qquad \text{converges in } \mathbb{R} \text{ for a.e. } x\ [d\mu]. \qquad (**)$$

By redefining the functions f_k as 0 on a set of μ measure 0, we may assume that the series $(**)$ converges pointwise to a limit in $\mathbb{R}$ for every x. Consequently the series

$$\sum_{k=2}^{\infty} (f_k(x) - f_{k-1}(x))$$

is absolutely convergent for all x and must be convergent for all x. The partial sums for the series without the absolute value signs are $f_n(x) - f_1(x)$, and hence $f(x) = \lim f_n(x)$ exists in $\mathbb{R}$ for every x. For every n,

$$|f - f_n| \le \sum_{k=n+1}^{\infty} |f_k - f_{k-1}| \le g, \qquad (\dagger)$$

and we have seen that g^p is integrable. By dominated convergence, we conclude that $\lim_n \int_X |f - f_n|^p\,d\mu = \int_X \lim_n |f(x) - f_n(x)|^p\,d\mu(x) = 0$. In other words, $\lim_n \|f - f_n\|_p = 0$. Therefore f_n tends to f in $L^p(\mu)$.

Next suppose that $p = \infty$. Let $\{f_n\}$ be any Cauchy sequence in L^∞. For each m and n, let E_{mn} be the subset of X where $|f_m - f_n| > \|f_m - f_n\|_\infty$, and put $E = \bigcup_{m,n} E_{mn}$. This set has measure 0. Redefine all functions to be 0 on E. The sequence of redefined functions is then uniformly Cauchy, hence uniformly convergent to some function f, and then f_n tends to f in $L^\infty(X)$.

For any p, we have shown that the original Cauchy sequence $\{f_n\}$ has a convergent subsequence $\{f_{n_k}\}$ in L^p. Let f be the L^p limit of the subsequence. Given $\epsilon > 0$, choose N such that $n \ge m \ge N$ implies $\|f_n - f_m\|_p \le \epsilon$, and then choose K such that $\|f_{n_k} - f\|_p \le \epsilon$ for $k \ge K$. Fix $k \ge K$ with $n_k \ge N$. Taking $m = n_k$, we see that $\|f_n - f\|_p \le \|f_n - f_{n_k}\|_p + \|f_{n_k} - f\|_p \le 2\epsilon$ whenever $n \ge n_k$. Thus $\{f_n\}$ converges to f. This completes the proof of the theorem. $\square$

In Section 8 we introduced integration of functions with values in $\mathbb{R}^m$ or $\mathbb{C}^m$. The definitions of L^1, L^2, and L^∞ may be extended to include such functions, and we write $L^1(X, \mathbb{C}^m)$, for example, to indicate that the functions in question take values in $\mathbb{C}^m$. In the definitions any expression $|f(x)|$ or $|f|$ that arises in the definition and refers to absolute value in the scalar-valued case is now to be understood as referring to the norm on the vector space where the functions take their values. The vector-valued L^1, L^2, and L^∞ spaces are further normed linear spaces, and one readily checks that Theorem 5.59 with the above proof applies to them because the range spaces are complete.

The triangle inequality for a pseudo normed linear space says that the norm of the sum of two elements is less than or equal to the sum of the norms, and of course the inequality instantly extends to a sum of any finite number of elements. But what about an integral of elements? In the case that the linear space is one of the precursor spaces "V" for L^1, L^2, or L^∞, the setting is that of functions of two variables. One of the variables corresponds to the measure space under study, and the other corresponds to the indexing set for the integral of the norms. Thus we could, if we wanted, force the situation into the mold of vector-valued functions whose values are in a space of functions. But it is not necessary to do so, and we do not. Here is the theorem.

Theorem 5.60 (Minkowski's inequality for integrals). Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and put $p = 1, 2$, or ∞. If f is measurable on $X \times Y$, then

$$\Big\| \int_X f(x, y)\, d\mu(x) \Big\|_{p, d\nu(y)} \le \int_X \|f(x, y)\|_{p, d\nu(y)}\, d\mu(x)$$

in the following sense: The integrand on the right side is measurable. If the integral on the right is finite, then for almost every y $[d\nu]$ the integral on the left is defined; when it is redefined to be 0 for the exceptional y's, then the formula holds.

REMARK. An extension of this theorem to values of p other than $1, 2, \infty$ will be given in Chapter IX, and that result will have the same name.

PROOF. The right side of the integral formula is unchanged if we replace f by $|f|$, and thus we may assume that $f \ge 0$ without loss of generality. If $p = 1$, then the formula for $f \ge 0$ reads

$$\textstyle\int_Y \Big[\int_X f(x, y)\, d\mu(x) \Big] d\nu(y) \overset{?}{\le} \int_X \Big[\int_Y f(x, y)\, d\nu(y) \Big] d\mu(x).$$

In fact, equality holds, and the result just amounts to Fubini's Theorem (Theorem 5.47).

Let $p = 2$. We have

$$\|f(x, y)\|^2_{2,d\nu(y)} = \int_Y |f(x, y)|^2 \, d\nu(y),$$

and this is measurable by Fubini's Theorem. Hence $\|f(x, y)\|_{2,d\nu(y)}$ is measurable. The idea for proving the inequality in the statement of the theorem is to imitate the argument that derives the triangle inequality for L^2 from the Schwarz inequality. That earlier argument is

$$\|g + h\|_2^2 = \|g\|_2^2 + 2\,\mathrm{Re}(g, h) + \|h\|_2^2 \le \|g\|_2^2 + 2\|g\|_2\|h\|_2 + \|h\|_2^2.$$

The adapted argument is

$$\begin{aligned}
\Big\| \textstyle\int_X f(x, y)\, d\mu(x)\Big\|^2_{2,d\nu(y)} &= \textstyle\int_Y \int_{x\in X} f(x, y)\, d\mu(x) \int_{x'\in X} \overline{f(x', y)}\, d\mu(x')\, d\nu(y) \\
&= \textstyle\int_{X\times X} \big[\int_Y f(x, y)\overline{f(x', y)} d\nu(y)\big]\, d\mu(x)\, d\mu(x') \\
&\le \textstyle\int_{X\times X} \|f(x, y)\|_{2,d\nu(y)} \|f(x', y)\|_{2,d\nu(y)}\, d\mu(x)\, d\mu(x') \\
&= \big[\textstyle\int_X \|f(x, y)\|_{2,d\nu(y)}\, d\mu(x)\big]^2,
\end{aligned}$$

the second and third lines following from Fubini's Theorem and the Schwarz inequality.

Let $p = \infty$. This is the hard case of the proof. We proceed in three steps. The first step is to prove the asserted measurability of $\|f(x, y)\|_{\infty,d\nu(y)}$, and we do so by first handling simple functions and then passing to the limit. If $s = \sum_{n=1}^N c_n I_{E_n}$ is the canonical expansion of a simple function $s \ge 0$ on $X \times Y$ and if x is fixed, then $\|s(x, y)\|_{\infty,d\nu(y)} = \max\big\{c_n \bigm| \nu((E_n)_x) > 0\big\}$. In other words, if k_n is the indicator function of the set $\big\{x \in X \bigm| \nu((E_n)_x) > 0\big\}$, then $s = \max\{c_1 k_1, \dots, c_N k_N\}$. Each function $c_n k_n$ is measurable by Lemma 5.44, and the pointwise maximum s is measurable by Corollary 5.9. Returning to our function $f \ge 0$, we use Proposition 5.11 to choose an increasing sequence $\{s_n\}$ of nonnegative simple functions with pointwise limit f. We prove that $\|s_n(x, y)\|_{\infty,d\nu(y)}$ increases to $\|f(x, y)\|_{\infty,d\nu(y)}$ for each x, and then the measurability follows from Corollary 5.10. Since x is fixed in this step, let us drop it and consider an increasing sequence $\{s_n\}$ of nonnegative measurable functions on Y with limit f on Y; we are to show that $\|f\|_\infty = \lim \|s_n\|_\infty$. The numbers $\|s_n\|_\infty$ are monotone increasing and are $\le \|f\|_\infty$. Thus $\lim \|s_n\|_\infty \le \|f\|_\infty$. Arguing by contradiction, suppose that equality fails and that $\lim \|s_n\|_\infty \le M < M+\epsilon < \|f\|_\infty$. Then $\big\{y \bigm| s_n(y) \ge M+\epsilon\big\}$ has measure 0 for every n, and so does $\bigcup_n \big\{y \bigm| s_n(y) \ge M + \epsilon\big\}$, by complete additivity. On the other hand, $\big\{y \bigm| f(y) > M + \epsilon\big\}$ is a subset of this union, and it has positive measure since $M + \epsilon < \|f\|_\infty$. Thus we have a contradiction and conclude that $\lim \|s_n\|_\infty = \|f\|_\infty$. Consequently $\|f(x, y)\|_{\infty,d\nu(y)}$ is measurable, as asserted.

The second step is to prove that any measurable function $F \geq 0$ on Y has $\|F\|_\infty = \sup_g \big|\int_Y F g\, d\nu\big|$, where the supremum is taken over all $g \geq 0$ with $\|g\|_1 \leq 1$. Certainly any such g has $\big|\int_Y F g\, d\nu\big| \leq \|F\|_\infty \int_Y g\, d\nu \leq \|F\|_\infty$, and therefore $\sup_g \big|\int_Y F g\, d\nu\big| \leq \|F\|_\infty$. For the reverse inequality, let I_E be the indicator function of a set of finite positive measure, and put $g = \nu(E)^{-1} I_E$. Then $\int_Y F g\, d\nu = \nu(E)^{-1} \int_E F\, d\nu \geq \inf_E(F)$. If m is less than $\|F\|_\infty$, then the set E where F is $\geq m$ has positive measure, and the inequality reads $m \leq \int_Y F g\, d\nu$ for the associated g. Hence $m \leq \sup_g \int_Y F g\, d\nu$. Taking the supremum of such m's, we obtain $\|F\|_\infty \leq \sup_g \big|\int_Y F g\, d\nu\big|$, and the reverse inequality is proved.

The third step is to use the previous two steps to prove the inequality in the statement of the theorem for $f \geq 0$. Let g be any nonnegative function on Y with $\int_Y g\, d\nu \leq 1$. Then Fubini's Theorem, the result of the first step above, and the result in the easy direction of the second step above give

$$\begin{aligned}\int_Y g(y)\big[\int_X f(x,y)\, d\mu(x)\big]\, d\nu(y) &= \int_X \big[\int_Y f(x,y) g(y)\, d\nu(y)\big]\, d\mu(x)\\ &\leq \int_X \big[\|f(x,y)\|_{\infty, d\nu(y)}\big]\, d\mu(x).\end{aligned}$$

Taking the supremum over g and using the result in the hard direction of the second step, we obtain the inequality in the statement of the theorem. $\square$

10. Arc Length and Lebesgue Integration

Section III.11 took up the topic of arc length for simple arcs $\gamma : [a,b] \to \mathbb{R}^n$. For any partition $P = \{t_j\}_{j=0}^m$ of $[a,b]$, we wrote $\ell(\gamma(P))$ for the sum of the lengths of the line segments connecting the consecutive points $\gamma(t_j)$, namely $\ell(\gamma(P)) = \sum_{j=1}^m |\gamma(t_j) - \gamma(t_{j-1})|$, and we defined

$$\ell(\gamma) = \sup_P \ell(\gamma(P)),$$

the supremum being taken over all partitions P of $[a,b]$. We called γ **rectifiable** if $\ell(\gamma)$ is finite.

In practice the simple arcs of most interest are the ones for which γ is of class C^1 on (a,b). We saw in Section III.11 on the one hand that not every simple arc of this kind is rectifiable but that the simple arcs of this kind with $|\gamma'|$ bounded are indeed rectifiable. We saw on the other hand that the theory omits vital examples if we consider only simple arcs in this class for which $|\gamma'|$ is bounded.

To handle this gap, we studied those simple arcs that are "tamely behaved" in the sense of being of class C^1 on (a,b) and having the property that near each endpoint, each entry of γ' is either bounded below or bounded above. These arcs were sufficient for our purposes. They were all rectifiable, and we derived the formula

$$\ell(\gamma) = \lim_{\substack{a'\downarrow a,\ b'\uparrow b,\\ a<a'<b'<b}} \int_{a'}^{b'} |\gamma'(t)|\,dt.$$

Armed with Lebesgue integration, we can sort out these matters and see exactly which simple arcs under study were rectifiable. The answer is as follows.

Proposition 5.61. A simple arc $\gamma : \gamma : [a, b] \to \mathbb{R}^n$ that is of class C^1 on (a, b) is rectifiable if and only is $|\gamma'|$ is Lebesgue integrable on $[a, b]$ with respect to Lebesgue measure m, and then

$$\ell(\gamma) = \int_{[a,b]} |\gamma'|\,dm.$$

PROOF. Whenever $a < a' < b' < b$, Theorem 3.42 and Example 3 of Section 2 show that

$$\ell(\gamma_{[a',b']}) = \textstyle\int_{a'}^{b'} |\gamma'(t)|\,dt = \int_{[a',b')} |\gamma'|\,dm.$$

Since the Lebesgue integral is a completely additive set function (Theorem 3.19) and since the one-point sets $\{a\}$ and $\{b\}$ have Lebesgue measure 0, we obtain

$$\lim_{\substack{a'\downarrow a,\ b'\uparrow b,\\ a<a'<b'<b}} \ell(\gamma_{[a',b']}) = \textstyle\int_{(a,b)} |\gamma'|\,dm = \int_{[a,b]} |\gamma'|\,dm.$$

Proposition 3.38 shows that the limit on the left side equals $\ell(\gamma)$ if γ is rectifiable, i.e., if $\ell(\gamma) < \infty$, and the proof will be complete if we show that $\int_{[a,b]} |\gamma'|\,dm = \infty$ when $\ell(\gamma) = \infty$.

Arguing by contradiction, suppose that $\ell(\gamma) = \infty$ and that $\int_{[a,b]} |\gamma'|\,dm = C < \infty$. Let M be an upper bound for $|\gamma(t)|$ for $a \le t \le b$. Because $\ell(\gamma) = \infty$, we can choose a partition P with $\ell(\gamma(P)) \ge C + 4M + 1$, say $P = \{t_j\}_{j=0}^m$. Without loss of generality, we may assume that the points t_j are distinct. Put $a' = t_1$ and $b' = t_{m-1}$. Then we have

$$\ell(\gamma(P)) = |\gamma(a') - \gamma(a)| + \sum_{j=2}^{m-1} |\gamma(t_j) - \gamma(t_{j-1})| + |\gamma(b) - \gamma(b')|.$$

The first and third terms on the right side are each $\le 2M$, and the middle term is $\gamma_{[a',b']}(P')$ for the partition $P' = \{t_j\}_{j=1}^{m-1}$ of $[a', b']$. Thus

$$C + 4M + 1 \le \ell(\gamma(P)) \le 4M + \ell(\gamma_{[a',b']}(P') \le 4M + \ell(\gamma_{[a',b']}).$$

The formula of the proposition has been proved for $\gamma_{[a',b']}$, and thus $C + 1 \le \ell(\gamma_{[a',b']}) = \int_{[a',b']} |\gamma'|\,dm \le \int_{[a,b]} |\gamma|\,dm = C$. Since C has been assumed finite, this inequality is a contradiction, and the result follows. $\square$

Corollary 5.62. If a simple arc $\gamma : [a, b] \to \mathbb{R}^n$ with γ of class C^1 on (a, b) is tamely behaved, then $|\gamma'|$ is integrable on $[a, b]$.

PROOF. This is immediate from Theorem 3.42 and Proposition 5.61. $\square$

REMARK. It is instructive to verify Corollary 5.62 by direct calculation. We omit the details.

11. Problems

1. Let X be a finite set of $n > 0$ elements.
   (a) If $\mathcal{A}$ is an algebra of subsets, what are the possible numbers of sets in $\mathcal{A}$?
   (b) Show that symmetric difference $A \,\Delta\, B = (A - B) \cup (B - A)$ is an abelian group operation on the set of all subsets of X and that every nontrivial element has order 2.
   (c) If $\mathcal{B}$ is a class of subsets containing $\varnothing$ and X and closed under symmetric difference, what are the possible numbers of sets in $\mathcal{B}$?
   (d) Prove or disprove: The class of sets in (c) is necessarily an algebra of sets.
   (e) Show that intersection and symmetric difference satisfy the distributive law $A \cap (B \,\Delta\, C) = (A \cap B) \,\Delta\, (A \cap C)$.
2. Exhibit a completely additive set function ρ on a σ-algebra and two sets A and B such that $\rho(A) < 0$ and $\rho(B) < 0$ but $\rho(A \cup B) > 0$.
3. Let $\{E_n\}$ be a sequence of subsets of X, and put
$$A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \quad \text{and} \quad B = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k.$$
   Prove that the indicator functions of E_k, A, and B satisfy
$$I_A = \limsup_n I_{E_n} \quad \text{and} \quad I_B = \liminf_n I_{E_n}.$$
4. Suppose that μ is a finite measure defined on a σ-algebra and $\{E_n\}$ is a sequence of measurable sets with
$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E_k.$$
   Call the set on the two sides of this equation E. Prove that $\lim_n \mu(E_n)$ exists and equals $\mu(E)$.
5. Let X be the set of rational numbers, and let $\mathcal{R}$ be the ring of all finite disjoint unions of bounded intervals in X, with or without endpoints. For each set E in $\mathcal{R}$, let $\mu(E)$ be its length.
   (a) Show that μ is nonnegative additive.
   (b) Show that μ is not completely additive.

6. Prove that if E is a Lebesgue measurable subset of $[0, 1]$ of Lebesgue measure 0, then the complement of E is dense in $[0, 1]$.

7. Let μ be a measure defined on a σ-algebra. Prove that if the complement of every set of measure $+\infty$ is of finite measure, then $\sup_{\mu(A)<+\infty} \mu(A)$ is finite and there is a set B with $\mu(B) = \sup_{\mu(A)<+\infty} \mu(A)$.

8. If f is a measurable function, prove that $f^{-1}(E)$ is measurable whenever E is a Borel subset of the real line.

9. For the measure space $(X, \mathcal{A}, \mu)$ in which X is the positive integers, $\mathcal{A}$ consists of all subsets of X, and μ is the counting measure, the theory of Lebesgue integration becomes a theory of infinite series. Restate Fatou's Lemma and the Dominated Convergence Theorem in this context.

10. Suppose on a finite measure space that $\{f_n\}$ is a sequence of real-valued integrable functions tending uniformly to f. Prove that $\lim_n \int_X f_n \, d\mu = \int_X f \, d\mu$.

11. This problem involves a Cantor set C in $[0, 1]$ built using fractions r_n as in Section II.9.
    (a) Show that C has Lebesgue measure $\prod_{n=1}^{\infty} (1 - r_n)$.
    (b) Prove that the indicator function I_C is discontinuous at every point of C and only there. Thus the set of discontinuities of I_C is not of measure 0 if $\prod_{n=1}^{\infty} (1 - r_n) > 0$.
    (c) Show that if the result of redefining I_C on a set of Lebesgue measure 0 is a function f, then the only possible points of continuity of f are those where f is 0.
    (d) Conclude that there exists a bounded Lebesgue measurable function on $[0, 1]$ that is not Riemann integrable and cannot be redefined on a set of measure 0 so as to be Riemann integrable.

12. Let $(X, \mathcal{A}, \mu)$ be any measure space, and let $(X, \overline{\mathcal{A}}, \overline{\mu})$ be its completion. Prove that if f is a function measurable with respect to $\overline{\mathcal{A}}$, then f can be redefined on a set of $\overline{\mu}$-measure 0 so as to be measurable with respect to $\mathcal{A}$.

13. Let X be an uncountable set, and let $\mathcal{A}$ be the set of all countable subsets of X and their complements. Prove that the diagonal $\{(x, x) \mid x \in X\}$ is not a member of the σ-algebra $\mathcal{A} \times \mathcal{A}$, the smallest σ-algebra containing all rectangles with sides in $\mathcal{A}$.

14. Let $(\mathbb{R}^1, \mathcal{B}, m)$ be the real line with Lebesgue measure on the Borel sets, and let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. If $f \geq 0$ is a measurable function on X, prove that the "region under the graph of f," defined by
$$R = \{(x, y) \mid 0 \leq y < f(x)\},$$
is a measurable subset of $X \times \mathbb{R}^1$ and that its measure relative to $\mu \times m$ is $\int_X f(x) \, d\mu(x)$.

15. Let $\mathcal{A}$ be a σ-algebra of subsets of a nonempty set X, let $F : \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_k} \to \mathbb{C}^N$ be continuous, and let $f_j : X \to \mathbb{C}^{n_j}$ be measurable with respect to $\mathcal{A}$ for $1 \le j \le k$. Prove that $x \mapsto F(f_1(x), \dots, f_k(x))$ is measurable with respect to $\mathcal{A}$.

16. This problem complements the proof in Theorem 5.59 that L^1 is a complete metric space. For $n \ge 1$, suppose that $0 < a_n < 1$ and $\sum_{n=1}^{\infty} a_n = +\infty$. Find a measure space $(X, \mathcal{A}, \mu)$ and a sequence of functions f_n with $\|f_n\|_1 = a_n$ and $\{f_n(x)\}$ convergent for no x.

17. **(Egoroff's Theorem)** Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Suppose that f_n and f are measurable functions with values in $\mathbb{R}$ such that $\lim f_n(x) = f(x)$ pointwise. The objective of this problem is to prove that $\lim f_n = f$ "almost uniformly." By considering the sets
$$E_{MN} = \{x \in X \mid |f_n(x) - f(x)| < 1/M \text{ for } n \ge N\}$$
for M fixed and N varying, prove that if $\epsilon > 0$ is given, then there exists a measurable subset E of X with $\mu(E) < \epsilon$ such that $\lim f_n(x) = f(x)$ uniformly for x in E^c.

18. (a) Derive the Dominated Convergence Theorem for a space of finite measure from Egoroff's Theorem (Problem 17) and Corollary 5.24.
    (b) Derive the Dominated Convergence Theorem for a space of infinite measure from the Dominated Convergence Theorem for a space of finite measure.

Problems 19–21 use Egoroff's Theorem (Problem 17) to show how close pointwise convergence is to L^1 convergence on a measure space $(X, \mathcal{A}, \mu)$ of finite measure. Theorem 5.59 shows that if a sequence converges in $L^1(X)$, then a subsequence converges almost everywhere. These problems address the converse direction in a way different from Problem 16. Suppose that f_n and f are integrable functions with values in $\mathbb{R}$ such that $\lim f_n(x) = f(x)$ pointwise.

19. Suppose that $f_n \ge 0$ for all n and that $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$. Prove that $\lim_n \int_E f_n \, d\mu = \int_E f \, d\mu$ for every measurable set E.

20. Suppose that $f_n \ge 0$ for all n and that $\lim \int_X f_n \, d\mu = \int_X f \, d\mu$. Use the previous problem and Egoroff's Theorem to prove that $\lim \int_X |f_n - f| \, d\mu = 0$.

21. A sequence $\{g_n\}$ of nonnegative integrable functions is called **uniformly integrable** if for any $\epsilon > 0$, there is an N such that $\int_{\{x \mid f_n(x) \ge N\}} g_n \, d\mu < \epsilon$ for all n. Suppose that the members of the given convergent sequence $\{f_n\}$ are nonnegative. Using Egoroff's Theorem in one direction and the previous problem in the converse direction, prove that $\lim_n \int_X f_n \, d\mu = \int_X f \, d\mu$ if and only if the f_n are uniformly integrable.

Problems 22–24 concern the extension of measures beyond what is given in Theorem 5.5 and Proposition 5.37. Let μ be a finite measure on a σ-algebra $\mathcal{A}$ of subsets of X, and define μ_* and μ^* on all subsets of X as in Lemma 5.32 and immediately after it. Let E be a subset of X that is not in $\mathcal{A}$, and let $\mathcal{B}$ be the smallest σ-algebra containing E and the members of $\mathcal{A}$.

22. Show that there exist two sets K and U in $\mathcal{A}$ such that $K \subseteq E \subseteq U$, $\mu_*(E) = \mu(K)$, and $\mu^*(E) = \mu(U)$. Show that K and U have the further properties that $U^c \subseteq E^c \subseteq K^c$, $\mu_*(E^c) = \mu(U^c)$, and $\mu^*(E^c) = \mu(K^c)$.
23. Show that the sets K and U of the previous problem satisfy $\mu_*(A \cap E) = \mu(A \cap K)$ and $\mu^*(A \cap E) = \mu(A \cap U)$ for every A in $\mathcal{A}$.
24. Fix t in [0, 1]. Show that the set function σ defined for A and B in $\mathcal{A}$ by
$$\begin{aligned}&\sigma[(A \cap E) \cup (B \cap E^c)] \\ &\qquad = t\mu_*(A \cap E) + (1-t)\mu^*(A \cap E) + t\mu^*(B \cap E^c) + (1-t)\mu_*(B \cap E^c)\end{aligned}$$
is defined on all of $\mathcal{B}$, is a measure, agrees with μ on $\mathcal{A}$, and assigns measure $t\mu_*(E) + (1-t)\mu^*(E)$ to the set E.

Problems 25–33 concern a construction by "transfinite induction" of all sets in the smallest σ-algebra containing an algebra of sets. In particular, it describes how to obtain all Borel sets of the interval [0, 1] of the line from the elementary sets in that interval. Later problems in the set apply the construction in various ways. This set of problems makes use of partial orderings as described in Section A9 of Appendix A, but they do not use Zorn's Lemma. The set of **countable ordinals** is an uncountable partially ordered set Ω, under a partial ordering $\leq$, with the following properties:

(i) Ω has the property that $x \leq y$ and $y \leq x$ together imply $x = y$,
(ii) Ω is "totally ordered" in the sense that any x and y in the set have either $x \leq y$ or $y \leq x$,
(iii) Ω is "well ordered" in the sense that any nonempty subset has a least element,
(iv) for any x in Ω, the set of elements $\leq x$ is at most countable.

Take as known that such a set Ω exists.

25. Prove that any countable subset of Ω has a least upper bound.
26. This problem asks for a proof of the validity of **transfinite induction** as applied to Ω. Let 1 be the least element of Ω, and let "$<$" mean "$\leq$ but not $=$." Suppose that some $p(\omega)$ is specified for each ω in Ω. Suppose further that $p(1)$ is true and that if for each $\omega > 1$, $p(\omega')$ is true for all $\omega' < \omega$, then $p(\omega)$ is true. Prove that $p(\omega)$ is true for all ω in Ω.
27. Let X be a nonempty set, let $\mathcal{A}$ be an algebra of subsets of X, and let $\mathcal{B}$ be the smallest σ-algebra containing $\mathcal{A}$. This problem uses Ω to describe "constructively" $\mathcal{B}$ in terms of $\mathcal{A}$. We define by transfinite induction two successively larger classes of sets $\mathcal{U}_\alpha$ and $\mathcal{K}_\alpha$ for each countable ordinal $\alpha \geq 1$. Let $\mathcal{U}_1$ be the set of all countable increasing unions of members of $\mathcal{A}$, let $\mathcal{K}_\alpha$ for $\alpha \geq 1$ be the set of all countable decreasing intersections of members of $\mathcal{U}_\alpha$, and let $\mathcal{U}_\alpha$ for $\alpha > 1$ be the set of all countable increasing unions of members of previous $\mathcal{K}_\beta$'s.

(a) Prove at each stage α that $\mathcal{U}_\alpha$ and $\mathcal{K}_\alpha$ are both closed under finite unions and finite intersections.

(b) Prove that $\mathcal{B}$ is the union of all $\mathcal{K}_\alpha$ for α in Ω.

28. For the case that $\nu(X) < +\infty$, prove the uniqueness half of the Extension Theorem (Theorem 5.5) by using the transfinite construction of Problem 27. [Educational note: It is not known how to prove the existence half of the Extension Theorem in this "constructive" way.]

29. Prove the Monotone Class Lemma (Lemma 5.43) by making use of the transfinite construction of Problem 27.

30. Devise a transfinite construction of all finite-valued Borel measurable functions on $\mathbb{R}^1$ that starts from continuous functions and alternately allows pointwise increasing limits and pointwise decreasing limits. The construction is to be in the spirit of Problem 27. Show that all finite-valued Borel measurable functions are obtained in this way if the indexing is done with Ω.

31. This problem "counts" the number of Borel sets of the real line, using Problem 27. It uses the material on cardinality in Section A10 of Appendix A.

(a) Prove that

(i) Ω has the same cardinality as some subset of $\mathbb{R}$,

(ii) the set of all sequences of members of $\mathbb{R}$ has the same cardinality as $\mathbb{R}$,

(iii) if $A \subseteq B \subseteq C$ and if A and C have the same cardinality as $\mathbb{R}$, then so does B,

(iv) if a set A has the same cardinality as $\mathbb{R}$ and if for each α in A, B_α is a set with the same cardinality as $\mathbb{R}$, then $\bigcup_{\alpha\in A} B_\alpha$ has the same cardinality as $\mathbb{R}$.

(b) Deduce that the set of all Borel sets of $\mathbb{R}$ has the same cardinality as $\mathbb{R}$ itself.

32. The standard Cantor set C in [0, 1], built using fractions $r_n = 1/3$ as in Section II.9, is a Borel set of Lebesgue measure 0 by Problem 11. Prove that C has the same cardinality as $\mathbb{R}$. Conclude that the cardinality of the set of all Lebesgue measurable sets equals the cardinality of the set of all subsets of $\mathbb{R}$. [Educational note: From this and Problem 31 it follows that there exists a Lebesgue measurable set in [0, 1] that is not a Borel set.]

33. For the standard Cantor set C as in the previous problem, show that the indicator function $I_{C'}$ of any subset C' of C is continuous on C^c. Conclude that the cardinality of the set of Riemann integrable functions on [0, 1] equals the cardinality of the set of all subsets of $\mathbb{R}$. [Educational note: From this and Problems 30–31, it follows that there exists a Riemann integrable function on [0, 1] that is not Borel measurable.]

Problems 34–41 show how to produce nontrivial nonnegative additive set functions on the set of all subsets of an infinite set from Zorn's Lemma (Section A9 of Appendix A).

A **filter** $\mathcal{F}$ on a nonempty set X is a nonempty class of subsets of X such that

(i) if E is in $\mathcal{F}$ and $F \supseteq E$, then F is in $\mathcal{F}$, i.e., $\mathcal{F}$ is closed under the operation of forming supersets,

(ii) if E and F are in $\mathcal{F}$, so is $E \cap F$,

(iii) $\varnothing$ is not in $\mathcal{F}$.

An **ultrafilter** is a filter that is not properly contained in any larger filter.

34. Verify the following:
    (a) $\{X\}$ is a filter.
    (b) Any filter is closed under finite intersections.
    (c) A one-point set and all of its supersets form an ultrafilter. (Such an ultrafilter is called a **trivial ultrafilter**.)
    (d) If X is infinite, then the set $\mathcal{F}$ of all subsets whose complements are finite sets is a filter.

35. Use Zorn's Lemma to show that every filter is contained in some ultrafilter.

36. Show that if $\mathcal{C}$ is a nonempty class of subsets of X, then there is a filter containing $\mathcal{C}$ if and only if no finite intersection of members of $\mathcal{C}$ is empty.

37. Prove that a filter $\mathcal{F}$ is an ultrafilter if and only if $A \cup B$ in $\mathcal{F}$ implies that either A is in $\mathcal{F}$ or B is in $\mathcal{F}$.

38. Prove that a filter $\mathcal{F}$ is an ultrafilter if and only if for every $A \subseteq X$, either A is in $\mathcal{F}$ or A^c is in $\mathcal{F}$.

39. Prove that the nonzero additive set functions defined on the set of *all* subsets of a set X and having image $\{0, 1\}$ stand in one-one correspondence with the ultrafilters on X, the correspondence being that the sets in the ultrafilter are exactly the sets on which the set function is 1. Prove that the set function is a measure if and only if the corresponding ultrafilter is closed under countable intersections.

40. Let X be any infinite set. Prove that X has a nontrivial ultrafilter, hence that X has a nonnegative additive set function μ that assumes only the values 0 and 1 and is not a point mass.

41. Prove that the set $\mathbb{Z}^+$ of positive integers has no nontrivial ultrafilter closed under countable intersections, i.e., that the set function μ in the previous problem is not a measure.

Problems 42–43 concern a theory of integration in which complete additivity is dropped as an assumption. An example is given in Problems 39–41 of a nonnegative additive set function on the set of all subsets of an infinite set that is not completely additive. For the present set of problems, let X be a nonempty set, let $\mathcal{A}$ be a σ-algebra of subsets, and let μ be a nonnegative additive set function on $\mathcal{A}$ such that $\mu(X) < +\infty$. Imagine an integration theory for $\int_E f\,d\mu$ with the definitions just as in the case that μ is a measure. All the properties of the integral proved in the text before the Monotone Convergence Theorem would still be valid, except that the integral $\int_E f\,d\mu$ as a function of E would be merely additive, rather than completely additive, and hence we would have to drop Corollary 5.24 and the converse half of Corollary 5.23.

42. Let f be ≥ 0, and let s_n be the standard pointwise increasing sequence of simple functions with limit f, as in Proposition 5.11. Show that the convergence of s_n to f is uniform if f is bounded.

43. Use the result of the previous problem to show in this theory that $\int_E (f+g)\,d\mu = \int_E f\,d\mu + \int_E g\,d\mu$ if f and g are bounded and measurable.