Chapter 1

Integration on Measure Spaces

A measure space is a triple consisting of a non-empty set X, a σ-algebra (a subset of the power set of X), and a measure defined on the σ-algebra. Measurable functions can be defined whenever a σ-algebra is given on X, and the integration of a non-negative measurable function is possible whenever there is a measure on the σ-algebra. In the first three sections of this chapter we study measure spaces and measurable functions. As passing to infinity for functions is very common in analysis, it is necessary to consider functions taking infinity as their values, so the extended real number system is introduced in Section 2. Sections 4 and 5 are devoted to the integration theory. Starting with the integration of simple functions, we next integrate non-negative measurable functions and end up with arbitrary measurable functions. Properties of the integral and the behavior of sequences of functions under integration are studied too. In Section 6 we investigate the relations among four notions of convergence of functions. They are, namely, pointwise almost everywhere convergence, uniform convergence, L^1-convergence and convergence in measure.

1.1 Measurable Functions

Throughout these notes X always denotes a non-empty set. A *σ-algebra* $\mathcal{M}$ is a subset of the power set $\mathcal{P}_X$ of X satisfying (i) $X \in \mathcal{M}$, (ii) $A' \in \mathcal{M}$ whenever $A \in \mathcal{M}$, and (iii) $\displaystyle\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}$, whenever $A_j \in \mathcal{M}, j \geq 1$. Here A' denotes the complement of A in X. From (i)–(iii) we immediately deduce the following properties

- $\phi \in \mathcal{M}$;
- $B \backslash A \in \mathcal{M}$ for $A, B \in \mathcal{M}$, and

- $\bigcap_{j=1}^{\infty} A_j \in \mathcal{M}$ for $A_j \in \mathcal{M}, \quad j \geq 1$.

Consequently, a σ-algebra is closed under taking complement, and countable union and intersection. For any given X, there is at least one σ-algebra, namely, $\mathcal{P}_X$, the power set itself. Observe that the intersection of any collection of σ-algebras is still a σ-algebra. Thus for every subset $\mathcal{S}$ of $\mathcal{P}_X$, there is a minimal σ-algebra containing S called the *σ-algebra generated by S*. Taking $X = \mathbb{R}^n$, the n-dimensional Euclidean space with the topology induced by the Euclidean norm and $\mathcal{S}$ the collection of all open sets, the σ-algebra generated by $\mathcal{S}$ is called the *Borel σ-algebra* and its elements are called *Borel sets*. We will study Borel sets in Chapter 2.

With a σ-algebra $\mathcal{M}$ on X, one can talk about measurable functions. A real-valued function f is called *measurable* or, more precisely, *measurable with respect to $\mathcal{M}$* or *$\mathcal{M}$-measurable*, if $f^{-1}(G) \in \mathcal{M}$ for every open set G in $\mathbb{R}$. Recall that a set G is open if for every $x \in G$, there exists an open interval I containing x that is contained in G. Therefore, every open set can be written as a union of open intervals. In fact, one can show that the union can be taken to be a countable one. (Prove it.) Let $G = \bigcup_n I_n$ where $I_n, n \geq 1$, are open intervals. Using $f^{-1}(\bigcup_n I_n) = \bigcup_n f^{-1}(I_n)$ and the fact that a σ-algebra is closed under countable unions, we see that f is measurable if and only if $f^{-1}(a,b) \in \mathcal{M}$ for all $a, b, a < b$. In fact, we have

Proposition 1.1. *f is measurable if and only if one of (a)–(d) holds:*

(a) $f^{-1}(a, \infty) \in \mathcal{M}, \quad \forall a \in \mathbb{R}$,

(b) $f^{-1}[a, \infty) \in \mathcal{M}, \quad \forall a \in \mathbb{R}$,

(c) $f^{-1}(-\infty, a) \in \mathcal{M}, \quad \forall a \in \mathbb{R}$,

(d) $f^{-1}(-\infty, a] \in \mathcal{M}, \quad \forall a \in \mathbb{R}$.

Proof. (a) First claim $f^{-1}[a, \infty) \in \mathcal{M}$, for all $a \in \mathbb{R}$. Indeed, we have

$$f^{-1}[a, \infty) = \bigcap_{n=1}^{\infty} f^{-1}(a - 1/n, \infty) \qquad \text{(check!)}$$

As $f^{-1}(a - 1/n, \infty)$ is measurable for all n, $f^{-1}[a, \infty) \in \mathcal{M}$ is measurable. Using

$$f^{-1}(a, b) = f^{-1}(a, \infty) \backslash f^{-1}[b, \infty) \in \mathcal{M},$$

$f^{-1}(a, b)$ is measurable. The proofs of (b)–(d) are similar. □

Next we examine how measurability is preserved feasibly under composition of functions, algebraic operations, and in the process of passing to limits.

Proposition 1.2. *Let $\Phi : \mathbb{R} \to \mathbb{R}$ be continuous. For every measurable f, $\Phi \circ f$ is measurable.*

Recall that Φ is continuous means for each $x \in \mathbb{R}$ and $\{x_n\}$ such that $x_n \to x$, $f(x_n) \to f(x)$. It can be shown that this is equivalent to, $f^{-1}(G)$ is open for every open G.

Proof. For every open set G, $\Phi^{-1}(G)$ is open in $\mathbb{R}$. From $(\Phi \circ f)^{-1}(G) = f^{-1}(\Phi^{-1}(G)) \in \mathcal{M}$, so $\Phi \circ f$ is measurable. □

By choosing the continuous functions $\Phi(z) = z^2$ in this proposition, we deduce that f^2 is measurable provided f is measurable. By choosing other suitable Φ's, we conclude that f_+, f_-, and $|f|$ are measurable when f is measurable. We point out that the assumption $\Phi : \mathbb{R} \to \mathbb{R}$ can be relaxed to $\Phi : V \to \mathbb{R}$ where V is an open set in $\mathbb{R}$ and the same conclusion still holds. It is because $\Phi^{-1}(G)$ is now open in V. However, as V is open, it is also open in $\mathbb{R}$.

Proposition 1.3. *(a) All measurable functions form a vector space.*
(b) fg is measurable whenever f and g are measurable.
(c) f/g is measurable provided f and g are measurable and $g(x) \neq 0$ for every x.

Proof. (a) Clearly, a constant multiple of a measurable function is measurable. It suffices to show that $f + g$ is measurable for measurable f and g, but this follows from the formula

$$(f+g)^{-1}(a,\infty) = \bigcup_{\substack{t+s>a \\ t,s\in\mathbb{Q}}} f^{-1}(t,\infty) \cap g^{-1}(s,\infty) \in \mathcal{M}.$$

To verify this formula, let's consider a point x satisfying $(f+g)(x) > a$. We can always choose two rational numbers t and s such that $f(x) > t, g(x) > s$ and $t + s > a$. It follows that $x \in f^{-1}(t,\infty) \bigcap g^{-1}(s,\infty)$, and we have one side inclusion. The other side inclusion is immediate.
(b) From

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

we conclude that fg is measurable from (a).
(c) It suffices to show that $1/g$ is measurable. This follows from the remark right after the proof of Proposition 1.2. We simply take $\Phi(z) = 1/z$. An alternate, elementary proof runs as follows: For $a, b > 0$, we have $(1/g)^{-1}(a,b) = g^{-1}(1/b, 1/a)$ is measurable. A similar result holds when $a, b < 0$. When $a < 0 < b$, it suffices to observe that

$$\left(\frac{1}{g}\right)^{-1}(a,b) = \left(\frac{1}{g}\right)^{-1}(a,0) \bigcup \left(\frac{1}{g}\right)^{-1}(0,b)$$

and $(1/g)^{-1}(a,0) = \bigcup_n (1/g)^{-1}(a,-1/n)$. □

Proposition 1.4. *Let f_k, $k \geq 1$, be measurable. Then*

$$\sup_k f_k(x), \quad \inf_k f_k(x), \quad \overline{\lim_{k\to\infty}} f_k(x) \text{ and } \underline{\lim_{k\to\infty}} f_k(x)$$

are all measurable provided they are finite everywhere.

Proof. Letting

$$g(x) = \sup_k f_k(x),$$

we have

$$g^{-1}(a,\infty) = \bigcup_{k=1}^{\infty} f_k^{-1}(a,\infty) \in \mathcal{M},$$

so g is measurable. Similarly, $\inf_k f_k$ is measurable. On the other hand, by definition

$$\overline{\lim_{k\to\infty}} f_k(x) = \inf_k \sup_{j\geq k} f_j(x),$$

and

$$\underline{\lim_{k\to\infty}} f_k(x) = \sup_k \inf_{j\geq k} f_j(x),$$

so they are also measurable.

□

1.2 Extened Real Numbers

In real analysis we need to pass to limit all the time. It will be convenient to take infinity as the value of functions. For this reason, we need to extend the real number system to accommodate the infinity. The *extended real number system* $\overline{\mathbb{R}}$ is obtained from $\mathbb{R}$ by adding ∞ and $-\infty$, that is, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$. It carries a metric structure given by, for every $x, y \in \mathbb{R}$,

$$\begin{aligned}
d(x,y) &= |\tan^{-1} x - \tan^{-1} y|, \\
d(\infty,x) = d(x,\infty) &= \left|\frac{\pi}{2} - \tan^{-1} x\right|, \\
d(-\infty,x) = d(x,-\infty) &= \left|\tan^{-1} x + \frac{\pi}{2}\right|, \\
d(\infty,-\infty) = d(-\infty,\infty) &= \pi.
\end{aligned}$$

The induced topology of this metric structure on $\mathbb{R}$ coincides with the usual topology. Next, we extend the algebraic structure a little bit by setting

$$\begin{aligned}
&x + \infty = \infty + x = \infty, \\
&x - \infty = -\infty + x = -\infty, \\
&x \cdot \infty = \infty, \quad \forall x \in (0, \infty], \\
&x \cdot \infty = -\infty, \quad \forall x \in [-\infty, 0), \\
&0 \cdot \infty = 0, \quad 0 \cdot (-\infty) = 0, \\
&x \cdot (-\infty) = -\infty, \quad \forall x \in (0, \infty], \\
&x \cdot (-\infty) = \infty, \quad \forall x \in [-\infty, 0).
\end{aligned}$$

Note that $\infty - \infty$ is undefined. By a routine argument, one can show that the associative, distributive and commutative laws are valid as long as both sides of the formulas are well-defined. Moreover, $(x, y) \to x + y$ is continuous from $\overline{\mathbb{R}}/\{(\infty, -\infty), (-\infty, \infty)\}$ to $\overline{\mathbb{R}}$, and $(x, y) \to xy$ is continuous from $\overline{\mathbb{R}} \setminus \{(\pm\infty, 0), (0, \pm\infty)$ to $\overline{\mathbb{R}}$.

From now on the notation $\overline{\mathbb{R}}$ stands for the extended real numbers with the topological and algebraic structure described above.

We note that in calculus, a sequence $\{x_n\}$ tends to infinity means for each $M > 0$, there exists some n_0 such that

$$x_n > M, \qquad \forall n \geq n_0$$

and the notation $x_n \to \infty$ is used. In $\overline{\mathbb{R}}$, $x_n \to \infty$ allows another interpretation; now ∞ is a point in $\overline{\mathbb{R}}$. It means

$$d(x_n, \infty) \to 0 \quad \text{as } n \to \infty.$$

It is easy to see that both interpretations are consistent.

An extended real-valued function f is *measurable* if $f^{-1}(G) \in \mathcal{M}$ for every open G open $\overline{\mathbb{R}}$. As every open set in $\overline{\mathbb{R}}$ can be written as the countable union of intervals of the form (a, b), $(a, \infty]$, $[-\infty, a)$, $a, b \in \mathbb{R}$, f is measurable if and only if $f^{-1}(a, b)$, $f^{-1}(a, \infty]$, $f^{-1}[-\infty, a)$ are in $\mathcal{M}$. In fact, it suffices to check $f^{-1}(a, b)$, $f^{-1}(\infty)$, $f^{-1}(-\infty) \in \mathcal{M}$. Corresponding to Propositions 1.2 and 1.4, we have

Proposition 1.2'. Let $\Phi : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ be continuous. For every measurable, extended real-valued f, $\Phi \circ f$ is measurable.

Proposition 1.4'. Let $f_k, k \geq 1$, measurable, extended real-valued functions. Then

$$\sup_k f_k(x), \quad \inf_k f_k(x), \quad \overline{\lim_{k \to \infty}} f_k(x), \text{ and } \underline{\lim_{k \to \infty}} f_k(x)$$

are measurable, extended real-valued functions.

As $\infty - \infty$ does not make sense, extended real-valued functions do not form a vector space.

1.3 Measure Space

Let $\mathcal{M}$ be a σ-algebra on X. A *measure* μ is a function from $\mathcal{M}$ to $[0, \infty]$ satisfying

(i) $\mu(\phi) = 0$ and

(ii) (Countable Additivity)

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k),$$

if $A_k \in \mathcal{M}$, $k \geq 1$, are mutually disjoint.

From (i) and (ii) one deduces, for every $A, B \in \mathcal{M}$,

- $\mu(B) = \mu(A) + \mu(B \backslash A)$ provided $A \subset B$.
- $\mu(A) \leq \mu(B)$, $\forall A \subset B$; and
- (Countable Subadditivity)

$$\mu(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mu(A_j), \quad A_j \in \mathcal{M},\ j \geq 1.$$

We recall the proof. Let $B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \bigcup A_2), \cdots$. Then B_j's are disjoint, $B_j \subset A_j$, and $\bigcup_j B_j = \bigcup_j A_j$. By countable additivity,

$$\mu(\bigcup_j A_j) = \mu(\bigcup_j B_j) = \sum_j \mu(B_j) \leq \sum_j \mu(A_j).$$

The triple $(X, \mathcal{M}, \mu)$ is called a *measure space*.

Example 1.1. For any subset E in X, define $c(E)$ to be the number of elements in E. Set $c(E)$ to be ∞ if E is an infinite set. Then $(X, \mathcal{P}_X, c)$ is called the *counting measure* on X.

Example 1.2. Let B_N be the collection of all sequences of the form $a_1 a_2 \cdots a_N$ where $a_j \in \{0, 1\}, j = 1, \cdots, N$. This is a finite set consisting of 2^N elements. For each singleton $\{x\}$ in the power set of B_N we assign $\mu(\{x\}) = 1/2^N$. The measure space $(B_N, \mathcal{P}_{B_N}, \mu)$ is called the *N-Bernoulli space*. In general, a measure space $(X, \mathcal{M}, \mu)$ satisfying $\mu(X) = 1$ a *probability measure*. Each element in B_N

represents the outcome of N many times of tossing a coin where 0 and 1 indicate a head and a tail respectively.

In general, a measure space $(X, \mathcal{M}, \mu)$ satisfying $\mu(X) = 1$ is called a *probability measure.* The Bernoulli spaces are probability spaces. A measure space is the setting for probability theory.

Example 1.3. Fix a point a in X and define $\delta_a(E) = 1$ or 0 according to whether $a \in E$ or not. The triple $(X, \mathcal{P}_X, \delta_a)$ is called the *Dirac measure* at a.

Example 1.4. We have learned that the Lebesgue measure $\mathcal{L}^1$ is a measure on $\mathcal{M}$, the σ-algebra containing all Lebesgue measurable sets in $\mathbb{R}$, which satisfies $\mathcal{L}^1\big((a, b)\big) = b - a$ for $a \le b$. It is known that $\mathcal{M}$ is a proper subset of $\mathcal{P}_{\mathbb{R}}$ which strictly contains all Borel sets. Remember the existence of non-measurable sets !

We have the following basic result.

Proposition 1.5. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

(a) Let A_k, $k \ge 1$, be an ascending sequence of measurable sets. We have

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \lim_{n\to\infty} \mu(A_n).$$

(b) Let A_k, $k \ge 1$, be a descending sequence of measurable sets. We have

$$\mu\Big(\bigcap_{k=1}^{\infty} A_k\Big) = \lim_{n\to\infty} \mu(A_n),$$

provided $\mu(A_1) < \infty$.

We call a sequence of sets A_k *ascending* (resp. *descending*) if $A_k \subset A_{k+1}$ (resp. $A_{k+1} \subset A_k$) for all k.

Proof. (a) Let $B_1 = A_1$, $B_2 = A_2 \backslash A_1$, $B_3 = A_3 \backslash A_1 \cup A_2$, $\cdots$ etc. Then $\{B_k\}$ is mutually disjoint and $\bigcup_{k=1}^{n} B_k = \bigcup_{k=1}^{n} A_k$, $n \leq \infty$. By countable additivity,

$$\begin{aligned}
\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) &= \mu\Big(\bigcup_{k=1}^{\infty} B_k\Big) \\
&= \sum_{k=1}^{\infty} \mu(B_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) \\
&= \lim_{n\to\infty} \mu\Big(\bigcup_{k=1}^{n} B_k\Big) \\
&= \lim_{n\to\infty} \mu(A_n).
\end{aligned}$$

(b) Let $C_k = A_1 \setminus A_k$, $k \geq 2$. Then $\{C_k\}$ is ascending and by (a),

$$\begin{aligned}
\mu\Big(\bigcup_{k=1}^{\infty} C_k\Big) &= \lim_{n\to\infty} \mu(C_n) \\
&= \lim_{n\to\infty} \mu(A_1 \setminus A_n) \\
&= \lim_{n\to\infty} \big(\mu(A_1) - \mu(A_n)\big) \\
&= \mu(A_1) - \lim_{n\to\infty} \mu(A_n).
\end{aligned}$$

As

$$\begin{aligned}
\bigcup_{k=1}^{\infty} C_k &= \bigcup_{k=1}^{\infty} (A_1 \bigcap A_k') \\
&= A_1 \bigcap \bigcup_{k=1}^{\infty} A_k' \\
&= A_1 \setminus \bigcap_{k=1}^{\infty} A_k \ ,
\end{aligned}$$

$$\begin{aligned}
\mu(\bigcup_{k=1}^{\infty} C_k) &= \mu\big(A_1 \setminus \bigcap_{k=1}^{\infty} A_k\big) \\
&= \mu(A_1) - \mu\big(\bigcap_{k=1}^{\infty} A_k\big) \ ,
\end{aligned}$$

and the result follows as $\mu(A_1) < \infty$. $\square$

We point out that the assumption $\mu(A_1) < \infty$ in Proposition 1.5 (b) is needed. To see this let $X = \mathbb{R}$ and consider the descending intervals $A_k = [k, \infty), k \geq 1$. Then $\bigcap_{k=1}^{\infty} A_k = \phi$ and $\mathcal{L}^1\big(\bigcap_{k=1}^{\infty} A_k\big) = 0$. However, on the other hand, $\mathcal{L}^1(A_k) = \infty$ for all k.

1.4 Integration on Measure Space

Let $(X, \mathcal{M}, \mu)$ be a measure space. A *simple function* is a measurable, real-valued function in X whose range is a finite set. It can be expressed in the form

$$s(x) = \sum_{j=1}^{N} \alpha_j \chi_{A_j}(x)$$

where $\alpha_1 < \alpha_2 < \cdots < \alpha_N$ and $A_j = \{x \in X : s(x) = \alpha_j\}, j \geq 1$, are mutually disjoint. It is clear that such representation of a simple function is unique. In general, a function of the form $s(x) = \sum_{j=1}^{N} \gamma_j \chi_{E_j}$ where $\gamma_j \in \mathbb{R}$ and $E_j, j = 1, \cdots, N$, are not necessarily mutually disjoint measurable sets, is a simple function since its range is a finite set. It can always be expressed in the form $\sum_{j=1}^{M} \alpha_j \chi_{F_j}$ where α_j are distinct and in ascending order and F_j are mutually disjoint and measurable. We may call it the "standard form" of s.

Simple functions are not to be confused with step functions. Recall that a step function is a function defined on $[a, b]$ which can be expressed in the form $\sum_k \alpha_k \chi_{I_k}$, where the sum is finite and I_k's are bounded intervals. A step function is always a simple functions (for the Lebesgue measure), but not every simple function is a step function. Moreover, step functions can only be defined on $\mathbb{R}$.

The following result demonstrates the importance of simple functions.

Theorem 1.6. *Let f be a non-negative, measurable extended real-valued function in $(X, \mathcal{M}, \mu)$. There exists a sequence of non-negative simple functions $\{s_k\}$, $k \geq 1$, such that for all x,*

(a) $s_k(x) \leq s_{k+1}(x)$, $\forall k \geq 1$; *and*

(b) $s_k(x) \to f(x), \quad$ *as* $\; k \to \infty$.

Proof. We define, for each k, a function φ_k on $[0, \infty]$ as follows. Divide $[0, \infty)$ into subintervals $[j/2^k, (j+1)/2^k), j = 0, 1, \ldots, k2^k - 1$, and $[k, \infty)$ and define $\varphi_k(t) = j/2^k$ for $t \in [j/2^k, (j+1)/2^k)$ and $\varphi_k(t) = k$ for $t \geq k$. Then $\varphi_k \leq \varphi_{k+1}$ for all k. Next we define $s_k(x) = \varphi_k(f(x))$, $\forall x \in X$. From the definition of φ_k we know that

$$s_k(x) \leq f(x) \leq s_k(x) + \frac{1}{2^k}, \quad \text{if } f(x) \in [0, k)$$

and $f(x) \geq s_k(x)$ if $f(x) \geq k$. It follows that $s_k(x) \to f(x)$, $\forall x$. The last thing we need to check is that s_k is measurable. But this follows from the fact that φ_k is a step function where the inverse image of an interval under φ_k is an interval. Check it. $\square$

In this section we discuss how to integrate a measurable function over a measurable space $(X, \mathcal{M}, \mu)$. We will accomplish this in three steps, first for for simple functions, next for non-negative measurable functions and finally for measurable functions.

In the following we sometimes call a function measurable in X without referring to the σ-algebra and the measure when it is easily identified in the context.

We briefly describe the ideas behind the abstract integration theory in light of Lebesgue and Riemann integrals. In fact, for a non-negative function f defined on $[0,1]$, say, its integral over $[0,1]$ is the area between the x-axis and the graph of f bounded by the two vertical lines $x = a, b$. Consider first the function $f_1(x) = 1$. Clearly its area is equal to 1. Next consider $f_2(x)$ which is equal to 0 when $x = 1/2^k, k \geq 1$, and 1 otherwise. The area of f_2 exists and is equal to 1. To see why the area is equal to 1, for each small $\varepsilon > 0$, consider the rectangles $R_k = [1/2^k - \varepsilon/2^{k+1}, 1/2^k + \varepsilon/2^{k+1}] \times [0,1], k \geq 1$. The total sum of area of these rectangles are bounded by $\sum_k \varepsilon/2^k = \varepsilon$ (Note that some R_k may overlap.) As f_1 is equal to 1 outside $[1/2^k - \varepsilon/2^{k+1}, [1/2^k + \varepsilon/2^{k+1}]$, the area of f_2 is bounded between $1 - \varepsilon$ and 1. As ε can be arbitrarily small, the area of f_2 is equal to 1. Now, let $\{x_k\}$ be a sequence in $[0,1]$ and let f_3 be the function which is equal to 0 at x_k and 0 elsewhere. The same argument above shows that the area of f_3 should be equal to 1. As all rational numbers form a countable set, we can take $\{x_k\}$ to be the set of all rational numbers in $[0,1]$. Then $f_3(x)$ becomes the function $f_4(x)$ which is equal to 0 when x is rational and to 1 when x is irrational. Although we have argued and been convinced that the area of f_4 is equal to 1, it is well-known that f_4 is not Riemann integral. While our argument is quite reasonable, there must be something wrong with Riemann integral. In fact, the defect of Riemann integral lies on its definition where it is required the Riemann sums converge to the same number as their lengths shrink to 0. Since there are rational and irrational numbers in any interval, the Riemann sums for any partition could be 1 or 0 depending on the tags, and thus do not converge to a definite value. This undesirable situation can be remedied by considering more general sums rather than the Riemann sums. Indeed, it is more reasonable to define the area of f over $[0,1]$ to be

$$\sup\Big\{ \sum_j \alpha_j |E_j| : \ s(x) = \sum_j \alpha_j \chi_{E_j}(x) \leq f(x), \forall x \in [0,1], E_j \subset [0,1] \Big\},$$

where $|E|$ denotes the "length" of E. Note that in Riemann integral one essentially takes E_j's to be disjoint intervals. In the example above, the area of f_4 is attained by taking $s(x) = 1 \times \chi_{[0,1]\setminus\mathbb{Q}} + 0 \times \chi_{\mathbb{Q}\cap[0,1]}$. Thus integration theory is built upon measure theory and measure theory was developed in an endeavor to understand the meaning of length for any set E.

Let s be a non-negative simple function given in standard form $\sum_{j=1}^{N} \alpha_j \chi_{A_j}$. We define the *integral* of s over the measurable set E to be, in notation,

$$\int_E s\, d\mu = \sum_{j=1}^{N} \alpha_j \mu(E \cap A_j).$$

Although the values of a simple function are finite, $\mu(E \cap A_j)$ could be infinite. From this definition we see that the value of the integral lies in $[0, \infty]$.

Recall that a function of the form $\sum_{j=1}^{N} \beta_j \chi_{E_j}$ where E_j's are measurable is a simple function. We first show that its integral can be expressed in terms of β_j and E_j without requiring them to be distinct or mutually disjoint.

Proposition 1.7. *Let $s = \sum_{j=1}^{N} \beta_j \chi_{E_j}$ be a non-negative simple function. Then*

$$\int_E s\, d\mu = \sum_{j=1}^{N} \beta_j \mu(E \cap E_j), \qquad \forall E \in \mathcal{M}.$$

Consequently,

$$\int_E (s+t) d\mu = \int_E s d\mu + \int_E t d\mu,$$

where t is a non-negative simple function.

Proof. First of all, we observe that this formula holds when β_j's are not necessarily distinct and E_j's are mutually disjoint.

To treat the general case, we can decompose $\{E_j\}$ into a mutually disjoint family of non-empty sets $\{F_i\}$ such that each F_i is contained in some E_j and each E_j is the union of those F_i's lying inside E_j. Indeed, each F_i is of the form $A_1 \cap A_2 \cdots \cap A_N$ where A_k is either E_k or E_k' (but excluding $E_1' \cap E_2' \cdots \cap E_N'$). There are at most $2^n - 1$ of them. We have

$$\begin{aligned} s &= \sum \beta_j \chi_{E_j} \\ &= \sum_j \beta_j \sum_{F_i \subset E_j} \chi_{F_i} \\ &= \sum_i \Big(\sum_{j,\ F_i \subset E_j} \beta_j \Big) \chi_{F_i} \\ &= \sum_i \gamma_i \chi_{F_i}, \quad \gamma_i \equiv \sum_{j,\ F_i \subset E_j} \beta_j . \end{aligned}$$

As $\{F_i\}$ are mutually disjoint but γ_i's are not necessarily distinct. By the remark beginning in this proof,

$$\begin{aligned}\int_E s\,d\mu &= \sum \gamma_i \mu(E\cap F_i)\\ &= \sum_i \sum_{j,\ \text{all}\ F_i\subset E_j} \beta_j \mu(E\cap F_i)\\ &= \sum_j \beta_j \sum_{i,\ F_i\subset E_j} \mu(E\cap F_i)\\ &= \sum_j \beta_j \mu(E\cap E_j),\end{aligned}$$

and the result follows. □

Next, for a measurable function $f : X \to [0,\infty]$ and a measurable set E, we define its *integral over* E to be, in notation,

$$\int_E f\,d\mu = \sup\left\{\int_E s\,d\mu : s \le f \text{ a.e., } s \text{ is a non-negative simple function}\right\}.$$

Here $s \le f$ a.e. means that s is less than or equal to f almost everywhere, in other words, there exists a measurable set N of measure zero such that $s \le f$ in $X \setminus N$. The definition is consistent with the principle that integration over a set of measure zero is equal to zero even though the integrand reaches infinity somewhere. Alternatively one could define the integral to be

$$\int_E f\,d\mu = \sup\left\{\int_E s\,d\mu : s \le f,\ s \text{ is a non-negative simple function}\right\}.$$

It is not hard to see that both definitions are the same. Indeed, let $s, s \le f$ a.e., we can produce another simple function $s_1, s_1 \le f$ everywhere, satisfying

$$\int_E s_1 d\mu = \int_E s d\mu\ .$$

Let $s = \sum_j \alpha_j \chi_{E_j}$. We set $A_j = \{x \in E_j :\ s(x) > f(x)\}$. Then $\cup_{j=1}^N A_j$ is a null set. The simple function $s_1 = \sum_j \alpha_j \chi_{E_j \setminus A_j}$ satisfies our requirement.

We point out that when f is a simple function, these definitions coincide with the integral of a simple function defined before.

Here are some elementary properties of the integral.

Proposition 1.8. *Let f and g be measurable functions from X to $[0,\infty]$. We have* *(a)* $\displaystyle\int_E f\,d\mu = \int_X f\chi_E\,d\mu, \quad \forall E \in \mathcal{M},$

(b) $\displaystyle\int_X g\,d\mu \ge \int_X f\,d\mu$ *if* $g \ge f$ *a.e.. When* $\displaystyle\int_X g\,d\mu$ *is finite, equality holds if and only if* $g = f$ *a.e.,*

(c) $\displaystyle\int_{E_1} f\,d\mu \le \int_{E_2} f\,d\mu$ *for measurable* $E_1, E_2,\ E_1 \subset E_2,$

(d) $\displaystyle c\int_X f\,d\mu = \int_X cf\,d\mu, \quad \forall c \in [0,\infty).$

In the following we drop the subscript X when the domain of integration is over the whole space.

Proof. (a) For any $s = \sum_k \alpha_k \chi_k \le f$ a.e., we have

$$\int_E s d\mu = \sum_k \alpha_k \mu(E \cap E_k) = \int_X s\chi_E d\mu,$$

and the conclusion follows from the definition of the integral.
(b) Obviously, when $f \le g$ a.e.,

$$\int f\,d\mu \le \int g\,d\mu.$$

Now, suppose that $g > f$ on some $A \in \mathcal{M}$, $\mu(A) > 0$. Letting $A_n = \big\{x \in A : g(x) > f(x) + 1/n\big\}$, we have $A = \bigcup_{n=1} A_n$. By Proposition 1.5(a) there exists some A_{n_0} with positive measure such that $g > f + 1/n_0$ on A_{n_0}. For any simple $s \le f$ a.e.,

$$g \ge s + \frac{1}{n_0}\chi_{A_{n_0}} \text{ a.e.},$$

so

$$\int g\,d\mu \ge \int s\,d\mu + \frac{1}{n_0}\mu(A_{n_0}),$$

by Proposition 1.7. (b) follows from this inequality after taking supremum over s.
(c) It follows from combining (a) and (b).
(d) Use the definition of the integral.

□

Next we have

Proposition 1.9 (Markov's Inequality). *Let f be a non-negative measurable function from X to $[0,\infty]$ with finite integral. Then*

$$\mu\big(\{x \in X: \ f(x) \geq M\}\big) \leq \frac{1}{M}\int_X f d\mu,$$

for all $M > 0$. Consequently

(a) Every integrable function is finite a.e..

(b) $f = 0$ a.e. if f is integrable and $\int f = 0$.

Proof. Let

$$A_M = \{x \in X: \ f(x) \geq M\} \ .$$

We have $f \geq f\chi_{A_M}$, so

$$\begin{aligned} \int f d\mu &\geq \int f\chi_{A_M} d\mu \\ &= \int_{A_M} f d\mu \\ &\geq \int_{A_M} M d\mu = M\mu(A_M) \ , \end{aligned}$$

from which the inequality follows. To prove (a), we note that $A_\infty \subset A_M$ for all finite $M > 0$. When $\int f d\mu$ is finite, we let $M \to \infty$ in the Markov's inequality to get

$$\begin{aligned} \mu(A_\infty) &\leq \lim_{M\to\infty} \mu(A_M) \\ &\leq \lim_{M\to\infty} \frac{1}{M}\int f d\mu \\ &= 0 \ . \end{aligned}$$

To prove (b), we note that

$$P = \{x \in X: \ f(x) > 0\} = \bigcup_{k=1}^{\infty} A_{A_{1/k}}.$$

But Markov's inequality tells us that $\mu(A_{1/k}) = 0$ for all k when the integral of f vanishes. It follows that

$$\mu(P) \leq \sum_{k=1}^{\infty} \mu(A_{1/k}) = 0 \ ,$$

so f is equal to zero a.e. .

□

Now we establish two basic results, namely, Lebsegue's monotone convergence theorem and Fatou's lemma.

Given a sequence of measurable functions which converges to some function almost everywhere, we would like to study when their integrals converge to the integral of the limit function. However, examples show that this is not always true. Here are some typical situations.

Example 1.5 (Mass leaking at peaks)**.** Take $\varphi_k = 0$ in $(1/k, 1]$, $\varphi_k(0) = k$ and linear between $(0, 1/k)$. Then $\varphi_k \to 0$ except at $x = 0$, so

$$\int \lim_{k\to\infty} \varphi_k \, d\mathcal{L}^1 = 0 \quad \text{but} \quad \int \varphi_k \, d\mathcal{L}^1 = 1, \ \forall k.$$

Example 1.6 (Mass vanishing at infinity)**.** Take $f_k = \chi_{[k,k+1]}$. Then $\lim_{k\to\infty} f_k(x) = 0$, so

$$\int \lim_{k\to\infty} f_k \, d\mathcal{L}^1 = 0 \quad \text{but} \quad \int f_k \, d\mathcal{L}^1 = 1, \ \forall k.$$

Example 1.7 (Mass dispersing away)**.** Take $g_k = k^{-1}\chi_{[0,k]}$, $k \geq 1$. Then $\lim_{k\to\infty} g_k(x) = 0$ for all x, so

$$\int \lim_{k\to\infty} f_k \, d\mathcal{L}^1 = 0 \quad \text{but} \quad \int g_k \, d\mathcal{L}^1 = 1, \ \forall k.$$

In light of these examples we need to impose further assumptions on the sequence in order to achieve this goal. The following theorem, making use of a monotonicity assumption, is one of the most frequently used results in integration theory.

Theorem 1.10 (Lebesgue's Monotone Convergence Theorem). *Let $f, f_k, k \geq 1$, be extended real-valued nonnegative measurable functions in the measure space $(X, \mathcal{M}, \mu)$. Suppose that there exists a measurable set N of zero measure such that f, f_k are non-negative and satisfy $f_k(x) \uparrow f(x)$ for $x \in X \setminus N$. Then*

$$\lim_{k\to\infty} \int f_k \, d\mu = \int f \, d\mu.$$

Proof. As $f \geq f_k$ a.e., by Proposition 1.8(b),

$$\int f \, d\mu \geq \int f_k \, d\mu, \quad \forall k,$$

which implies

$$\int f \, d\mu \geq \lim_{k\to\infty} \int f_k \, d\mu.$$

The limit on the right hand side exists because $\{f_k\}$ is increasing. On the other hand, fix some $\delta \in (0,1)$. For any $s \le f$ a.e., there exists a set of measure zero N_1 such that $s \le f$ in $X \setminus N_1$. Let $Y = X \setminus (N \cup N_1)$ and $E_k = \{x \in Y : f_k(x) \ge \delta s(x)\}$. Clearly, E_k is ascending. We claim that $\bigcup_{k=1}^{\infty} E_k = Y$. For, if $x \in Y$ and $f(x) = 0$, then $s(x) = 0$ and $x \in E_k$ for all k. If $\infty > f(x) > 0$, then $\delta s(x) \le \delta f(x) < f_k(x)$ for all large k. When $f(x) = \infty$, $f_k(x) \to \infty$, so $\delta s(x) < f_k(x)$ for all large k too.

Consequently, for these k,

$$\begin{aligned} \int f_k \, d\mu &\ge \int_{E_k} f_k \, d\mu \\ &\ge \delta \int_{E_k} s \, d\mu \\ &= \delta \sum_{j=1}^{N} \alpha_j \mu(E_k \cap A_j), \qquad \text{if } s = \sum \alpha_j \chi_{A_j}. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{k\to\infty} \int f_k \, d\mu &\ge \delta \sum_{j=1}^{N} \alpha_j \lim_{k\to\infty} \mu(E_k \cap A_j) \\ &= \delta \sum_{j=1}^{N} \alpha_j \mu(Y \cap A_j) \quad \text{(Proposition 1.5(a))} \\ &= \delta \sum_{j=1}^{N} \alpha_j \mu(A_j) \quad (\because \ \mu(N \cup N_1) = 0) \\ &= \delta \int s \, d\mu \ . \end{aligned}$$

Taking supremum over s,

$$\lim_{k\to\infty} \int f_k \, d\mu \ge \delta \int f \, d\mu$$

and

$$\lim_{k\to\infty} \int f_k \, d\mu \ge \int f \, d\mu$$

holds after letting $\delta \uparrow 1$.

□

The next result is also a frequently used one.

Theorem 1.11 (**Fatou's Lemma**). *Let $\{f_k\}$ be a sequence of extended real-*

valued measurable functions which are non-negative a.e.. Then

$$\int \varliminf_{k\to\infty} f_k \, d\mu \le \varliminf_{k\to\infty} \int f_k \, d\mu.$$

Proof. Set $g_k(x) = \inf_{j\ge k} f_j(x)$. Then each g_k is measurable and $\{g_k\}$ is an increasing sequence. The function g given by

$$g(x) \equiv \lim_{k\to\infty} g_k(x) = \sup_k g_k(x) = \varliminf_{k\to\infty} f_k(x) \ge 0$$

is measurable. By Lebesgue's monotone convergence theorem,

$$\begin{aligned}
\int \varliminf_{k\to\infty} f_k \, d\mu &\equiv \int g \, d\mu \\
&= \lim_{k\to\infty} \int g_k \, d\mu \\
&= \varliminf_{k\to\infty} \int g_k \, d\mu \\
&\le \varliminf_{k\to\infty} \int f_k \, d\mu \ .
\end{aligned}$$

□

Previous examples show that strict inequality could occur in Fatou's lemma.

Now we establish the linearity of the integral.

Proposition 1.12. *Let f and g be extended real-valued, nonnegative measurable functions in X and $\alpha, \beta \ge 0$. Then*

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

Proof. By Theorem 1.6, we can find sequences of simple functions $\{s_j\}$, $\{t_j\}$, $s_j \uparrow f$, $t_j \uparrow g$ in X. Thus $\alpha s_j + \beta t_j \uparrow f + g$ and

$$\begin{aligned}
\int (\alpha f + \beta g) \, d\mu &= \lim_{j\to\infty} \int (\alpha s_j + \beta t_j) \, d\mu \quad \text{(monotone convergence theorem)} \\
&= \lim_{j\to\infty} \left(\int \alpha s_j \, d\mu + \int \beta t_j \, d\mu \right) \quad \text{(Proposition 1.7)} \\
&= \int \alpha f d\mu + \int \beta g d\mu \\
&= \alpha \int f \, d\mu + \beta \int g \, d\mu. \quad \text{(Proposition 1.8)}
\end{aligned}$$

□

So far we have been considering non-negative functions. Now we come to functions which may change sign. For a measurable function f, its positive and negative parts f_+ and f_- are non-negative and measurable. The integrals

$$\int f_+ \, d\mu \quad \text{and} \quad \int f_- \, d\mu$$

both make sense. If either one of these integrals is finite, we define

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

An extended real-valued measurable function f in X is called *integrable* if both $\int f_+ \, d\mu$ and $\int f_- \, d\mu$ are finite. Using Proposition 1.12, when a function f is integrable, we have

$$\begin{aligned} \int |f| d\mu &= \int (f_+ + f_-) d\mu \\ &= \int f_+ d\mu + \int f_- d\mu < \infty, \end{aligned}$$

so $|f|$ is also integrable. It follows from Markov's inequality that every integrable function is finite a.e. .

Proposition 1.13. *Let f and g be integrable in X and α, $\beta \in \mathbb{R}$ such that $\alpha f + \beta g$ is well-defined in X. Then $\alpha f + \beta g$ is integrable and*

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

Moreover,

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

Since f and g are extended real-valued and integrable, they are finite almost everywhere. The combination $\alpha f + \beta g$ may not be defined over a null set. When this happens, we may assign ∞ (or any other fixed value) to $\alpha f + \beta g$ at all points in this set. Under this convention $\alpha f + \beta g$ is a measurable function defined everywhere and this proposition still holds.

Proof. Observe that $(f+g)_+$, $(f+g)_- \le |f+g| \le |f| + |g|$ so

$$\int (f+g)_\pm \, d\mu \le \int (|f| + |g|) \, d\mu = \int |f| \, d\mu + \int |g| \, d\mu < \infty.$$

We conclude that $f + g$ is integrable.

In the following we assume that f and g are finite everywhere. Otherwise we restrict to $X \setminus N$ where N is the null set consisting of all points either f or g becomes infinity. Using $f+g=(f+g)_+-(f+g)_-$ on one hand and $f+g=f_+-f_-+g_+-g_-$ on the other hand, we have

$$(f+g)_+ + f_- + g_- = (f+g)_- + f_+ + g_+ \quad \text{on } X.$$

Therefore,

$$\int (f+g)_+ \, d\mu + \int f_- \, d\mu + \int g_- \, d\mu = \int (f+g)_- \, d\mu + \int f_+ \, d\mu + \int g_+ \, d\mu$$

by Proposition 1.12. It follows that

$$\begin{aligned}
\int (f+g) \, d\mu &= \int (f+g)_+ d\mu - \int (f+g)_- d\mu \\
&= \int f_+ d\mu - \int f_- d\mu + \int g_+ d\mu - \int g_- d\mu \\
&= \int f \, d\mu + \int g \, d\mu.
\end{aligned}$$

Next, we need to show

$$\alpha \int f \, d\mu = \int \alpha f \, d\mu, \quad \forall \alpha \in \mathbb{R}.$$

When $\alpha \geq 0$, this is contained Proposition 1.12. To show it for $\alpha < 0$ it suffices to show

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

Indeed, we have

$$\int f \, d\mu + \int (-f) \, d\mu = \int \left[f + (-f) \right] d\mu = 0.$$

Finally, we have

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

□

The following result ensures when pointwise convergence implies convergence in the corresponding integrals. The required condition is the existence of a dominator for the sequence of functions under consideration. While Lebesgue's monotone convergence theorem and Fatou's lemma apply to non-negative functions, Lebesgue's dominated convergence theorem deals with functions that may change sign. It is more general.

Theorem 1.14 (Lebesgue's Dominated Convergence Theorem). *Let f, f_k, $k \geq 1$, be extended real-valued measurable in X satisfying $f_k \to f$ a.e. and $|f_k| \leq g$ a.e. for some integrable g. Then f is integrable and*

$$\lim_{k\to\infty} \int |f_k - f| \, d\mu = 0.$$

Proof. By assumption, the function $2g - |f_k - f|$ is non-negative a.e.. We can apply Fatou's lemma to get

$$\begin{aligned} \int 2g \, d\mu &= \int \varliminf_{k\to\infty} (2g - |f_k - f|) \, d\mu \\ &\leq \varliminf_{k\to\infty} \int (2g - |f_k - f|) \, d\mu \\ &= \int 2g \, d\mu - \varlimsup_{k\to\infty} \int |f_k - f| \, d\mu. \end{aligned}$$

As $\displaystyle\int g \, d\mu < \infty$, we can cancel it from both sides to get

$$\varlimsup_{k\to\infty} \int |f_k - f| \, d\mu \leq 0.$$

□

Note that $\displaystyle\int |f_k - f| \, d\mu \to 0$ implies $\displaystyle\int f_k \, d\mu \to \int f \, d\mu$.

Corollary 1.15 (Tonelli's Theorem). *Let $a_k, k \geq 1$, be extended real-valued measurable and*

$$\sum_{k=1}^{\infty} \int |a_k(x)| \, d\mu < \infty.$$

Then $\displaystyle\sum_{k=1}^{\infty} a_k$ is integrable and

$$\int \sum_{k=1}^{\infty} a_k \, d\mu = \sum_{k=1}^{\infty} \int a_k \, d\mu < \infty \ .$$

Proof. Set

$$g_k(x) = \sum_{j=1}^{k} |a_j(x)|\,, \quad g(x) = \sum_{j=1}^{\infty} |a_j(x)|\,,$$

$$h_k(x) = \sum_{j=1}^{k} a_j(x), \quad h(x) = \sum_{j=1}^{\infty} a_j(x).$$

As $g_k \uparrow g$, by the monotone convergence theorem

$$\int g\,d\mu = \int \lim_{k\to\infty} g_k\,d\mu = \sum_{j=1}^{\infty} \int |a_j(x)|\;d\mu < \infty,$$

whence g is integrable. Now the corollary follows from using g as a dominator for h_k and applying the dominated convergence theorem. □

More convergence theorems can be found in Chapter 4.

1.5 Convergence of Measurable Functions

For a sequence of measurable functions from the measure space $(X, \mathcal{M}, \mu)$ to $\mathbb{R}$ or $\overline{\mathbb{R}}$, the following four notions of convergence make sense.

- $\{f_k\}$ converges to f almost everywhere, that is, there exists a null set N such that $f_k(x) \to f(x)$ for all $x \in X \setminus N$ as $k \to \infty$.

- $\{f_k\}$ converges to f uniformly on a subset Y, that is, for each $\varepsilon > 0$, there exists some k_0 such that $|f_k(x) - f(x)| < \varepsilon$, for all $k \geq k_0$ and $x \in Y$. (Assuming that $f_k(x)$ and $f(x)$ do not taking ∞ or $-\infty$ simultaneously.)

- $\{f_k\}$ converges in f in L^1-*sense*, that is, $\displaystyle\lim_{k\to\infty} \int |f_k - f|\,d\mu = 0$.

- $\{f_k\}$ converges to f *in measure*, that is, for each $\rho > 0$,

  $$\lim_{k\to\infty} \mu\big(\,\{x \in X : |f_k(x) - f(x)| \geq \rho\}\,\big) = 0.$$

  It is implicitly assumed that $f_k(x) - f(x)$ is well-defined almost everywhere. Note that the set $\{x \in X : |f_k(x) - f(x)| \geq \rho\}$ is measurable.

In this section we discuss some relations among them.

Theorem 1.16 (**Egorov Theorem**). *Let $f, f_k, k \geq 1$, be extended real-valued measurable functions in X which are finite a.e.. Suppose that $\mu(X)$ is finite and $f_k \to f$ a.e. as $k \to \infty$. Then for each $\varepsilon > 0$, there exists a measurable A, $\mu(A) < \varepsilon$, such that $f_k \to f$ uniformly on $X \setminus A$ as $k \to \infty$.*

Proof. WLOG assume that f_k and f are finite everywhere. Set

$$A_j^i = \bigcup_{k \ge j} \Big\{ x : |f_k(x) - f(x)| \ge \frac{1}{2^i} \Big\}.$$

For each fixed i, $\{A_j^i\}$ is a descending family in $\mathcal{M}$. As $\mu(X) < \infty$, by Proposition 1.5 we have

$$\mu\Big(\bigcap_{j=1}^{\infty} A_j^i \Big) = 0.$$

We can find a large $J(i)$ such that

$$\mu\Big(A_{J(i)}^i \Big) < \frac{\varepsilon}{2^i}.$$

The set $A = \bigcup_{i=1}^{\infty} A_{J(i)}^i$, satisfies

$$\mu(A) \le \sum_{i=1}^{\infty} \mu\big(A_{J(i)}^i\big) < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$$

We claim that $f_k \to f$ uniformly on $X \setminus A$. For, if $x \notin A$, x does not belong to $A_{J(i)}^i$ for all i. In other words,

$$|f_k(x) - f(x)| < \frac{1}{2^i}, \quad \forall k \ge J(i).$$

Given $\varepsilon > 0$, we can find a large i such that $2^{-i} < \varepsilon$. Then

$$|f_k(x) - f(x)| < \varepsilon, \quad \forall x \in X \setminus A, \quad \forall k \ge J(i).$$

□

Consider $f_k(x) = \chi_{[k,k+1]}, k \ge 1$. Then $\{f_k\}$ converges pointwisely to 0 as $k \to \infty$, but cannot converge to 0 uniformly away from a set of finite measure. It shows that the condition $\mathcal{L}^1(X) < \infty$ cannot be removed.

Proposition 1.17. *If $\{f_k\}$ converges to f in measure, then it admits a subsequence $\{f_{k_j}\}$ converging to f a.e. as $k_j \to \infty$.*

Proof. As $\{f_k\}$ converges to f in measure, we can pick a subsequence $\{f_{k_j}\}$ such that

$$\mu\Big(\Big\{ x : \big|f_{k_j}(x) - f(x)\big| \ge \frac{1}{j} \Big\} \Big) < \frac{1}{2^j}.$$

Set

$$B_n = \bigcup_{j \ge n} \Big\{ |f_{k_j} - f| \ge \frac{1}{j} \Big\},$$

where $\{\,|f_{k_j} - f| \geq 1/j\}$ is the short form for $\{x \in X : |f_{k_j}(x) - f(x)| \geq 1/j\}$. Then $\{B_n\}$ is descending and

$$\mu(B_1) = \mu\Big(\bigcup_{j=1}^{\infty}\Big\{\,|f_{k_j} - f| \geq \frac{1}{j}\Big\}\Big)$$
$$\leq \sum_{j=1}^{\infty}\mu\Big(\Big\{\,|f_{k_j} - f| \geq \frac{1}{j}\Big\}\Big)$$
$$\leq \sum_{j=1}^{\infty}\frac{1}{2^j} = 1.$$

Applying Proposition 1.5, we have

$$\mu\Big(\bigcap_{n=1}^{\infty} B_n\Big) = \lim_{n\to\infty}\mu(B_n)$$
$$\leq \lim_{n\to\infty}\sum_{j=1}^{\infty}\frac{1}{2^{n+j}} = 0.$$

The set $N \equiv \bigcap_{n=1}^{\infty} B_n$ is of measure zero. We claim that $\{f_{k_j}\}$ converges to f on $X \setminus N$. Indeed, for $x \in X \setminus N$, $x \in \bigcup_{n=1}^{\infty} B_n'$. There exists some n_1 such that $x \in B_{n_1}'$. From the definition of B_{n_1},

$$|f_{k_j}(x) - f(x)| < \frac{1}{j}, \quad \forall j \geq n_1,$$

so $\{f_{k_j}(x)\}$ converges to $f(x)$ for each $x \in X \setminus N$.

□

One may also prove this result by using Borel-Cantelli lemma.

Proposition 1.18. *Let $f, f_k, k \geq 1$, be measurable functions from X to $\overline{\mathbb{R}}$ which are finite a.e.. Suppose that $\mu(X)$ is finite and $f_k \to f$ a.e. as $k \to \infty$. Then $\{f_k\}$ converges to f in measure.*

Proof. By Egorov theorem, for $\varepsilon > 0$, there exists some measurable $A, \mu(A) < \varepsilon$, such that $f_k \to f$ uniformly on $X \setminus A$. Therefore, for $\rho > 0$, we can find k_0 such that $|f_k(x) - f(x)| < \rho$ for all $k \geq k_0$ and $x \in X \setminus A$. It follows that $\{x : |f_k(x) - f(x)| \geq \rho,\ x \in X, \forall k \geq k_0\} \subset A$ and so $\mu\big(\{x : |f_k(x) - f(x)| \geq \rho,\ x \in X, \forall k \geq k_0\}\big) < \varepsilon$. □

Proposition 1.19. *If $\{f_k\}$ L^1-converges to f, then $\{f_k\}$ converges to f in measure.*

Proof. For $\rho > 0$, let $E_k = \{x \in X : |f_k(x) - f(x)| \geq \rho\}$. As in the proof of Markov's inequality,

$$\int |f_k - f| \, du \geq \int_{E_k} |f_k - f| \, d\mu \geq \rho\mu(E_k),$$

which implies

$$\mu(E_k) \leq \frac{1}{\rho} \int |f_k - f| d\mu \to 0, \qquad k \to \infty.$$

□

Combining this proposition with Proposition 1.17, we arrive at the following useful fact.

Proposition 1.20. *If $\{f_k\}$ L^1-converges to f, then it has a subsequence converging almost everywhere to f.*

It is a good exercise to prove this fact directly.

Comments on Chapter 1. The study of an integration theory more flexible and general than the Riemann integration led Lebsegue invent his integration theory. In this theory one inevitably needs to define the "length" of a set in $[a, b]$] and measure theory was conceived. Subsequently developments, especially the mathematical formulation of probability theory by Kolmogorov, prompted the separation of measure theory from the theory of integration (although they are closely related). In this chapter we discuss how to define integration whenever a measure space is given. With differences only in some details, we largely follow [R] in this chapter. It is amazing that an integration theory can be fully developed based on the simple setting of a measure space. All results in this chapter are of fundamental nature, and will be used again and again in the subsequent development. Lebesgue's monotone convergence theorem, Fatou's lemma and Lebesgue's dominated convergence theorem are the corner stones for integration theory. You should understand them well. By the way, in these notes we deduce Fatou's lemma from Lebesgue's monotone convergence theorem. Alternatively one can prove Fatou's lemma first and then use it to deduce Lebesgue's monotone convergence theorem, see, for instance, [HS].

In [EG], a measurable function is called integrable if either $\int f_+$ or $\int f_-$ is finite, and it is called summable if both $\int f_+$ and $\int f_-$ are finite. Thus summable in [EG] means integrable in [R] as well as in these notes. Be careful of this discrepancy.

We do not consider complex-valued functions in these notes. Extending results from real-valued to complex-valued functions are usually straightforward and could be established on site when one really needs it.

Finally, *from now on a measurable function always means an extended real-valued measurable function. A null set is always a measurable set with zero measure.*