Notes on Measure Theory and Integration

John Villavert

Contents

Preface

These notes grew out of my lecture notes partially used in a first-semester real analysis course I taught at the University of Texas, Rio Grande Valley. The objective of these notes is to provide a relatively quick and concise introduction of the essentials of abstract measure theory and integration with emphasis in the setting of Euclidean spaces. Special emphasis is placed on the study of the Lebesgue integral and on the class of Radon measures defined on $\mathbb{R}^n$. These notes also accompany the principle textbooks of Bartle [2], Folland [3] and Rudin [4]

We try our best to keep the discussion self-contained but it is suggested the reader have undertaken a semester of undergraduate analysis and have familiarity with metric spaces, continuity, differentiation, and the construction of the Riemann integral.

Review of the Riemann integral and its limitations

We give a brief construction of the Riemann integral and its basic properties, and describe some of the deficiencies of this notion of integration in modern theory and applications in the mathematical sciences. For simplicity, we shall consider functions $f : [a, b] \longrightarrow \mathbb{R}$ that are defined and bounded on some closed and bounded interval $I = [a, b]$. Some preparations are in order to define what it means for f to be Darboux (or equivalently Riemann) integrable, but we focus on the Darboux construction of the integral.

Given an interval $I = [a, b]$, we may decompose I into n subintervals $\{I_i\}_{i=1}^n$, where

$$a = t_0 < t_1 < t_2 < t_3 \ldots < t_{n-1} < t_n = b,$$

and $I_i = t_i - t_{i-1}$ for $i = 1, 2, \ldots, n$. The length of each subinterval I_i will be denoted by $\ell(I_i) = t_i - t_{i-1}$.

Each such collection of subintervals of I is called a **subdivision** of I, and we denote it by $\Delta(I)$ or simply just by Δ. Given a subdivision Δ, the boundedness of f ensures the following quantities are defined:

$$m = \inf_{x\in I} f(x),\ M = \sup_{x\in I} f(x),\ m_i = \inf_{x\in I_i} f(x) \text{ and } M_i = \sup_{x\in I} f(x)$$

for $i = 1, 2, 3, \ldots, n$. Then we define the **upper Darboux sum** $S^+(f, \Delta)$ and the **lower Darboux sum** $S_-(f, \Delta)$ with respect to f and Δ, respectively by

$$S^+(f, \Delta) = \sum_{i=1}^{n} M_i \ell(I_i) \quad \text{and} \quad S_-(f, \Delta) = \sum_{i=1}^{n} m_i \ell(I_i).$$

Prior to defining the (Darboux) integral of a bounded function $f : [a, b] \longrightarrow \mathbb{R}$, some preliminary definitions and intermediate results are needed. Firstly, we may obtain a new subdivision from each subdivision $\Delta(I)$ with endpoints $\{t_i\}_{i=1}^n$ by adding additional points. This leads to a new subdivision called a **refinement** of Δ with subintervals $I_1', I_2', \ldots, I_m'$ such that each new subinterval I_j' is contained in a unique subinterval of the original subdivision Δ.

Now, given any pair of subdivisions Δ_1 and Δ_2 of the interval I, we can construct a new subdivision which is a proper refinement of both subdivisions. More precisely, let Δ_1 be a subdivision of I with subintervals $I_1, I_2, \ldots, I_n$ and let Δ_2 be another with subintervals $J_1, J_2, \ldots, J_m$. If we take all endpoints of all subintervals of both subdivisions and arrange them in increasing order then relabel accordingly, we obtain a new subdivision Δ that defines a refinement of both Δ_1 and Δ_2. This subdivision Δ is called the **common refinement** of Δ_1 and Δ_2.

Proposition 0.1. *Let f be a bounded function on $I = [a, b]$ and suppose Δ is a subdivision of I. Then*

(a) $m(b - a) \leq S_-(f, \Delta) \leq S_+(f, \Delta) \leq M(b - a)$.

(b) If Δ' is a refinement of Δ, then $S_-(f, \Delta) \leq S_-(f, \Delta') \leq S^+(f, \Delta') \leq S^+(f, \Delta)$.

(c) For any pair of subdivisions Δ_1 and Δ_2 of I, $S_-(f, \Delta_1) \leq S^+(f, \Delta_2)$.

Proof. We let Δ be a subdivision of I with n subintervals $I_i = [t_{i-1}, t_i]$ for $i = 1, 2, \ldots, n$.
(a) Noting that $m \leq m_i \leq M_i \leq M$ for $i = 1, 2, \ldots, n$, simple calculations reveal

$$\begin{aligned} m(b - a) = m \sum_{i=1}^{n} \ell(I_i) \leq \sum_{i=1}^{n} m_i \ell(I_i) = S_-(f, \Delta) \leq \sum_{i=1}^{n} M_i \ell(I_i) = S^+(f, \Delta) \\ \leq M \sum_{i=1}^{n} \ell(I_i) = M(b - a). \end{aligned}$$

(b) Let Δ' be a refinement of Δ, and let $I_1', I_2', \ldots, I_m'$ be its subintervals. As Δ' is itself a subdivision, the fact that

$$S_-(f,\Delta') \leq S^+(f,\Delta') \tag{0.1}$$

is a consequence of part (a).

By definition of a refinement, each subinterval I'_i is uniquely contained in a subinterval I_{k_i} in Δ for some $1 \leq k_i \leq n$. Likewise, each subinterval I_i in Δ can be expressed as a finite union of subintervals $\{I'_{i_j}\}_{j=1}^{N_i}$ in Δ', i.e., $I_i = \cup_{j=1}^{N_i} I'_{i_j}$.

Noticing that $m_i \leq m'_{i_j} := \inf_{x\in I'_{i_j}} f(x)$, we see that

$$\begin{aligned} S_-(f,\Delta) = \sum_{i=1}^{n} m_i \ell(I_i) &\leq \sum_{i=1}^{n} m_i \big(\sum_{j=1}^{N_i} \ell(I'_{i_j})\big) \\ &\leq \sum_{i=1}^{n}\sum_{j=1}^{N_i} m'_{i_j}\ell(I'_{i_j}) \leq \sum_{i=1}^{m} m'_i \ell(I'_i) = S_-(f,\Delta'). \end{aligned} \tag{0.2}$$

In addition, since $M_{k_i} \geq M'_i = \sup_{x\in I'_i} f(x)$, we obtain

$$S^+(f,\Delta') = \sum_{i=1}^{m} M'_i \ell(I'_i) \leq \sum_{i=1}^{n} M_i \ell(I_i) = S^+(f,\Delta). \tag{0.3}$$

Hence, combining (0.1)-(0.3) leads to the desired conclusion.

(c) For any subdivisions Δ_1 and Δ_2 of the interval I, let Δ be the common refinement of this pair. Part (b) implies $S_-(f,\Delta_1) \leq S_-(f,\Delta)$, part (a) implies $S_-(f,\Delta) \leq S^+(f,\Delta)$ and part (b) again implies $S^+(f,\Delta) \leq S^+(f,\Delta_2)$. Therefore, $S_-(f,\Delta_1) \leq S_+(f,\Delta_2)$. $\square$

We are ready to define the (Darboux) integral of a bounded function. First, we consider the sets

$$E_-(f) = \{S_-(f,\Delta) \,|\, \Delta \text{ is any subdivision of } I\}$$

and

$$E^+(f) = \{S^+(f,\Delta) \,|\, \Delta \text{ is any subdivision of } I\}.$$

Thanks to Proposition 0.1, both $E_-(f)$ and $E^+(f)$ are non-empty bounded subsets of $\mathbb{R}$ and therefore $\sup E_-$ and $\inf E^+$ exist, where the supremum and infimum are taken over all possible subdivisions of I. Then, we define the **upper Darboux integral** of f on $I = [a,b]$ to be the quantity $\inf E^+$, and we denote it by

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

The **lower Darboux integral** of f on I is defined by $\sup E_-$, which we denote by

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

We say a bounded function f is **Darboux integrable** or just **integrable** on I if $\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx$ and this common value is denoted by $\int_a^b f(x)\,dx$.

A consequence of Proposition 0.1 is the following theorem.

Theorem 0.1. *If f is a bounded function on $I = [a, b]$ such that $m(b-a) \leq f(x) \leq M(b-a)$ for all $x \in I$, then*

$$m(b-a) \leq \underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx \leq M(b-a).$$

A necessary and sufficient condition for the integrability of a bounded function on $I = [a, b]$ is provided in the following theorem.

Theorem 0.2 (Integrability Criterion)**.** *Suppose that f is bounded on an interval $I = [a, b]$. Then f is integrable if and only if the following property holds: For every $\epsilon > 0$ there exists a subdivision Δ of I such that*

$$S^+(f, \Delta) - S_-(f, \Delta) < \epsilon. \tag{0.4}$$

Proof. Assuming (0.4) holds, for any choice of $\epsilon > 0$, the definition of upper and lower Darboux sums lead to

$$\overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx < \epsilon.$$

Therefore, by the so-called ϵ-principle, this implies

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

On the other hand, Theorem 0.1 shows

$$\overline{\int_a^b} f(x)\,dx \geq \underline{\int_a^b} f(x)\,dx,$$

which verifies the upper and lower Darboux sums coincide. Hence, f is integrable on I.

Conversely, we assume f is integrable on I and let $\epsilon > 0$ be given. By definition of $\overline{\int_a^b} f(x)\,dx$ and $\underline{\int_a^b} f(x)\,dx$, there exist (why?) subdivisions Δ_1 and Δ_2 such that

$$S^+(f, \Delta_1) < \overline{\int_a^b} f(x)\,dx + \epsilon/2 \quad \text{and} \quad S_-(f, \Delta_2) > \underline{\int_a^b} f(x)\,dx - \epsilon/2. \tag{0.5}$$

Letting Δ be the common refinement of Δ_1 and Δ_2, Proposition (0.1) leads to

$$S^+(f, \Delta) - S_-(f, \Delta) \leq S^+(f, \Delta_1) - S_-(f, \Delta_2).$$

And inserting the inequalities of (0.5) into this last inequality will give us condition (0.4). This completes the proof of the theorem. □

Important consequences of this theorem are as follows. First, as any function f that is continuous on a compact set $[a, b]$ is uniformly continuous, the Integrability Criterion Theorem applies and ensures f is integrable on $[a, b]$. The theorem applies to bounded and monotonic functions as well, and we leave the details of both claims for the reader to verify.

Corollary 0.1. *If f is continuous on $I = [a, b]$, then f is integrable on I.*

Corollary 0.2. *If f is bounded and monotone on $I = [a, b]$, then f is integrable on I.*

Let us examine the basic properties of the integral.

Proposition 0.2. *Let k be a non-zero constant, and suppose f is a bounded function on $I = [a, b]$ and $g = kf$ on I.*

If $k > 0$, then

(a) $\underline{\int_a^b} g(x)\,dx = k\underline{\int_a^b} f(x)\,dx,$ *(b)* $\overline{\int_a^b} g(x)\,dx = k\overline{\int_a^b} f(x)\,dx.$

If $k < 0$, then

(a) $\underline{\int_a^b} g(x)\,dx = k\overline{\int_a^b} f(x)\,dx,$ *(b)* $\overline{\int_a^b} g(x)\,dx = k\underline{\int_a^b} f(x)\,dx.$

In order to prove this proposition, we shall need the following lemma.

Lemma 0.1. *Consider the same assumptions outlined in Proposition 0.2. Then*

If $k > 0$, then

(a) $\inf_{x\in I} g(x) = k \inf_{x\in I} f(x),$ *(b)* $\sup_{x\in I} g(x) = k \sup_{x\in I} f(x).$

If $k < 0$, then

(a) $\inf_{x\in I} g(x) = k \sup_{x\in I} f(x),$ *(b)* $\sup_{x\in I} g(x) = k \inf_{x\in I} f(x).$

Proof. We shall only prove parts (b) for the case of $k > 0$, since the remaining cases follow similarly. Let $k > 0$. Since $f(x_0) \le \sup_{x\in I} f(x)$ for all $x_0 \in I$, we get $g(x_0) = kf(x_0) \le k\sup_{x\in I} f(x)$ and so

$$\sup_{x\in I} g(x) \le k \sup_{x\in I} f(x). \tag{0.6}$$

To show the reverse inequality, we let $\epsilon > 0$ be given. By definition of the supremum, we can find an element $x_0 \in I$ such that $f(x_0) > \sup_{x\in I} f(x) - \epsilon/k$. This implies that $\sup_{x\in I} g(x) \ge g(x_0) = kf(x_0) > \sup_{x\in I} f(x) - \epsilon$, that is

$$k \sup_{x\in I} f(x) \le \sup_{x\in I} g(x) + \epsilon \quad \text{for every } \epsilon > 0.$$

The ϵ-principle implies

$$k \sup_{x\in I} f(x) \le \sup_{x\in I} g(x). \tag{0.7}$$

The desired identity in (b) follows from (0.6) and (0.7). □

Proof of Proposition 0.2. We prove parts (b) for both cases in k, since the remaining cases are proved similarly. Let $k > 0$ and suppose Δ be any subdivision of I with subintervals $I_1, I_2, \ldots, I_n$. Thanks to Lemma 0.1, for each $i = 1, 2, \ldots, n$, we have

$$\tilde{m}_i := \inf_{x \in I_i} g(x) = m_i := k \inf_{x \in I_i} f(x) \quad \text{and} \quad \tilde{M}_i := \sup_{x \in I_i} g(x) = M_i := k \sup_{x \in I_i} f(x).$$

This leads to

$$S^+(g, \Delta) = \sum_{i=1}^n \tilde{M}_i \ell(I_i) = \sum_{i=1}^n k M_i \ell(I_i) = k \sum_{i=1}^n M_i \ell(I_i) = k S^+(f, \Delta).$$

By taking the infimum over all possible subdivisions Δ of I gives $\inf S^+(g, \Delta) = k \inf S^+(f, \Delta)$, that is,

$$\overline{\int_a^b} g(x)\, dx = k \overline{\int_a^b} f(x)\, dx.$$

Instead, let $k < 0$, the above argument still applies and leads to $S^+(g, \Delta) = k S_-(f, \Delta)$. By taking the supremum over all possible subdivisions Δ of I and noting $k < 0$, we obtain

$$\sup S_-(g, \Delta) = \sup k S_-(f, \Delta) = k \inf S_-(f, \Delta),$$

This is equivalent to

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

□

Proposition 0.3. *Suppose f and g are bounded functions on $I = [a, b]$ such that $f(x) \leq g(x)$ for all $x \in I$. Then*

$$\text{(a)} \underline{\int_a^b} f(x)\, dx \leq \underline{\int_a^b} g(x)\, dx \quad \textit{and} \quad \text{(b)} \overline{\int_a^b} f(x)\, dx \leq \overline{\int_a^b} g(x)\, dx.$$

Proof. We only prove part (a) as the proof of (b) is similar. Take Δ to be any subdivision of I with subintervals $I_1, I_2, \ldots, I_n$ and denote $\tilde{m}_i = \inf_{x \in I_i} g(x)$, $\tilde{M}_i = \sup_{x \in I_i} g(x)$, $m_i = \inf_{x \in I_i} f(x)$, and $M_i = \sup_{x \in I_i} f(x)$ Since $f \leq g$ on I, we have $m_i \leq \tilde{m}_i$ and $M_i \leq \tilde{M}_i$ for $i = 1, 2, \ldots, n$ and thus

$$\begin{aligned} S_-(f, \Delta) &= \sum_{i=1}^n m_i \ell(I_i) \\ &\leq \sum_{i=1}^n \tilde{m}_i \ell(I_i) = S_-(g, \Delta) \leq \sup S_-(g, \Delta) = \underline{\int_a^b} g(x)\, dx. \end{aligned}$$

This leads to

$$\underline{\int_a^b} f(x)\, dx = \sup S_-(f, \Delta) \leq \underline{\int_a^b} g(x)\, dx.$$

□

Proposition 0.4. *Suppose f and g are bounded functions on $I = [a, b]$ and set $h = f + g$ on I. Then*

(a) $\underline{\int_a^b} h(x)\,dx \geq \underline{\int_a^b} f(x)\,dx + \underline{\int_a^b} g(x)\,dx,$

(b) $\overline{\int_a^b} h(x)\,dx \leq \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx.$

Proof. We shall only prove part (b) as part (a) is similar. Let Δ be any subdivision of I with subintervals $I_1, I_2, \ldots, I_n$. For all $x \in I_i$, $h(x) \leq f(x) + g(x) \leq \sup_{x \in I_i} f(x) + \sup_{x \in I_i} g(x)$, therefore

$$M_i := \sup_{x \in I_i} h(x) \leq \sup_{x \in I_i} f(x) + \sup_{x \in I_i} g(x) =: M_{1,i} + M_{2,i}.$$

Then we calculate the upper Darboux sum

$$S^+(h, \Delta) = \sum_i^n M_i \ell(I_i) \leq \sum_i^n M_{1,i} \ell(I_i) + \sum_i^n M_{2,i} \ell(I_i) = S^+(f, \Delta) + S^+(g, \Delta).$$

Thus,

$$\overline{\int_a^b} h(x)\,dx = \inf S^+(h, \Delta) \leq S^+(h, \Delta) \leq S^+(f, \Delta) + S^+(g, \Delta). \tag{0.8}$$

Now suppose Δ_1 and Δ_2 are any two subdivisions with subintervals $I_1, I_2, \ldots, I_n$ and $J_1, J_2, \ldots, J_m$, respectively, and let Δ denote their common refinement. From (0.8) and Proposition 0.1 we get

$$\overline{\int_a^b} h(x)\,dx \leq S^+(f, \Delta) + S^+(g, \Delta) \leq S^+(f, \Delta_1) + S^+(g, \Delta_2).$$

If we fix Δ_1 and let Δ_2 vary over all possible subdivisions of I, we deduce

$$\overline{\int_a^b} h(x)\,dx \leq S^+(f, \Delta_1) + \inf S^+(g, \Delta_2) = S^+(f, \Delta_1) + \overline{\int_a^b} g(x)\,dx.$$

From this inequality, taking the infimum over all possible subdivisions Δ_1 reveals

$$\overline{\int_a^b} h(x)\,dx \leq \overline{\int_a^b} f(x)\,dx + \overline{\int_a^b} g(x)\,dx.$$

□

Immediate consequences of Propositions 0.2-0.4 is the following set of results for integrable functions.

Theorem 0.3. *Let k be any constant, and suppose f and g are bounded and integrable functions on $I = [a, b]$. Then kf and $f + g$ are integrable such that*

$$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx \quad \text{and} \quad \int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

Moreover, if $f(x) \le g(x)$ for all $x \in I$, then

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

With these properties for the Riemann integral established, we can derive another useful consequence of Theorem 0.2.

Corollary 0.3. *Suppose that f is a bounded function on $I = [a, b]$. If f is integrable on I, then so is $|f|$. Moreover,*

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

Proof. Assume f is integrable on I. We prove this in two steps.

Step 1: We show the integrability of f implies the integrability of $|f|$.

Let $\epsilon > 0$ be given, and suppose f is integrable on $I = [a, b]$. By Theorem 0.2, there exists a subdivision Δ with subintervals $I_1, I_2, \ldots, I_n$ such that

$$S^+(f, \Delta) - S_-(f, \Delta) < \epsilon. \tag{0.9}$$

As before, on each subinterval we define

$$\tilde{M}_i = \sup_{I_i} |f(x)| \quad \text{and} \quad \tilde{m}_i = \inf_{I_i} |f(x)|$$

and

$$M_i = \sup_{I_i} f(x) \quad \text{and} \quad m_i = \inf_{I_i} f(x).$$

We also observe that for any pair of points $z_1, z_2 \in I_i$, there holds

$$|f(z_1) - f(z_2)| \le M_i - m_i. \tag{0.10}$$

So for any given $\varepsilon > 0$, the definition of the supremum and infimum ensure there are points $z_1, z_2 \in I_i$ such that $\tilde{M}_i - \varepsilon/2 < |f(z_1)|$ and $\tilde{m}_i > |f(z_2)| - \varepsilon/2$. Combining this with (0.9), the triangle inequality and (0.10) leads us to

$$\tilde{M}_i - \tilde{m}_i \le |f(z_1)| - |f(z_2)| + \varepsilon \le |f(z_1) - f(z_2)| + \varepsilon \le M_i - m_i + \varepsilon.$$

Thanks to the ε-principle and (0.9), this implies the estimate

$$\tilde{M}_i - \tilde{m}_i \le M_i - m_i$$

and therefore

$$\begin{aligned} S^+(|f|,\Delta) - S_-(|f|,\Delta) &= \sum_{i=1}^{n} (\tilde{M}_i - \tilde{m}_i)\ell(I_i) \\ &\le \sum_{i=1}^{n} (M_i - m_i)\ell(I_i) \\ &= S^+(f,\Delta) - S_-(f,\Delta) \\ &< \epsilon. \end{aligned}$$

Since $\epsilon > 0$ was chosen arbitrarily, Theorem 0.2 applies to $|f|$, and we deduce $|f|$ is integrable on I.

Step 2: We establish the remaining inequality. Supposing $\int_a^b f(x)\,dx \ge 0$, the fact that $f \le |f|$ on I and using Theorem 0.3 give us

$$\Big|\int_a^b f(x)\,dx\Big| = \int_a^b f(x)\,dx \le \int_a^b |f(x)|\,dx.$$

Otherwise, if $\int_a^b f(x)\,dx < 0$, the fact that $-f \le |f|$ on I and the same reasoning above yields

$$\Big|\int_a^b f(x)\,dx\Big| = -\int_a^b f(x)\,dx = \int_a^b -f(x)\,dx \le \int_a^b |f(x)|\,dx.$$

This concludes the proof. □

Let us make some concluding remarks and highlight some of the limitations of the Riemann and Darboux integrals and motivate why a more general and robust theory of integration is needed.

As Theorem 0.2 suggests, we place rather stringent conditions for a function to be Riemann integrable, and let us now highlight some of its deficiencies. For instance, we consider the characteristic function $\varphi : [a,b] \longrightarrow \mathbb{R}$, where $\varphi(x) = \chi_{\mathbb{Q}\cap[a,b]}$; i.e., $\varphi(x) = 1$ if $x \in \mathbb{Q}\cap[a,b]$ and $\varphi \equiv 0$ otherwise in $\mathbb{Q}\cap[a,b]$. Obviously this function is bounded on $I = [a,b]$. However, it is not integrable on I, since if we take any subdivision Δ of $[a,b]$, $S^+(\varphi,\Delta) = 1$ and $S_-(\varphi,\Delta) = 0$ and thus

$$\overline{\int_a^b} \varphi(x)\,dx = 1 \ne 0 = \underline{\int_a^b} \varphi(x)\,dx.$$

The Riemann integral also behaves quite nicely but under ideal settings, e.g., it works well with functions that are at least continuous, and it remains closed under uniform convergence. More precisely, denoting by $C([a,b])$ the normed space of continuous real-valued functions on $[a,b]$ equipped with the uniform norm,

$$\|f\|_{sup} = \sup_{a\le x\le b} |f(x)| \ \text{ for } f \in C([a,b]),$$

then $C([a,b])$ is a Banach space. Therefore, if a sequence $\{f_n\}_{n=1}^{\infty}$ in $C([a,b])$ **converges uniformly** to a function f, i.e., $\lim_{n\to\infty} \|f_n - f\|_{sup} = 0$, two important properties hold. First, the limiting function f belongs to $C([a,b])$ and is therefore Riemann integrable on I. Secondly,

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx. \tag{0.11}$$

In other words, uniform convergence ensures we can exchange the limit with the Riemann integral. The big problem occurs when this notion of convergence is weakened to, say, pointwise convergence. By this we mean $\{f_n\}_{n=1}^{\infty}$ **converges pointwise** to f on $[a,b]$ if $\lim_{n\to\infty} f_n(x) = f(x)$ for all $a \leq x \leq b$. Under pointwise convergence, the integrability of the limiting function and (0.11) may no longer hold. Here is a nice example from [1] illustrating this. If $I = [a,b] = [0,1]$ and we define $\{r_n\}_{n=1}^{\infty}$ to be the increasing sequence of all rational numbers in $[0,1]$, then consider the functions $f_n : [0,1] \longrightarrow \mathbb{R}$ such that $f_n(x) = 1$ if $x \in \{r_1, r_2, \ldots, r_n\}$ and $f_n(x) = 0$ otherwise. It is straightforward to check for each $n \in \mathbb{N}$, f_n is Riemann integrable on $[0,1]$ with $\int_0^1 f_n(x)\,dx = 0$. Then $\{f_n\}_{n=1}^{\infty}$ converges pointwise to the above function φ in $[0,1]$, but φ is not even Riemann integrable.

These issues with the Riemann integral are significant in modern mathematical problems, and the examples above illustrate that a more unrestrictive and general framework of integration is needed—one that overcomes such limitations but still recovers the Riemann integral in the ideal settings. This is especially critical in analysis and applications, since problems often deal with various different (weaker) notions of a limit within spaces that are often classes of irregular (discontinuous) functions. The more general integral is the so-called Lebesgue integral, but we shall not confine our attention to the Lebesgue integral on just $\mathbb{R}$ or the n-dimensional Euclidean spaces. We shall study a more general framework starting with abstract measure theory and use it to build a general theory of integration.

CHAPTER 1

Introduction to Measures

This chapter introduces the basic notion of generating a positive measure on an abstract space.

1.1 Measures and basic concepts of measurability

Throughout, X shall denote a non-empty set.

1.1.1 Algebras and σ-algebras

Definition. *An* **algebra** *of sets on* X *is any non-empty collection* $\mathcal{A} \subset 2^X$ *that is closed under finite unions and complements, i.e.,* $E_1, E_2, \ldots E_n \in \mathcal{A}$ *implies that* $\cup_{i=1}^n E_i \in \mathcal{A}$*; and* $E^c \in \mathcal{A}$ *whenever* $E \in \mathcal{A}$.

Here, 2^X denotes the power set of X.

Definition. *A* σ-**algebra** *is just an algebra that is closed under countable unions.*

It is easy to verify the finite union or arbitrary intersection of any collection of σ-algebras on X is a σ-algebra. The latter observation ensures that given $\mathcal{E} \subset 2^X$, we have the following.

Lemma 1.1. *There exists a unique smallest* σ*-algebra denoted by* $\mathcal{M}(\mathcal{E})$ *containing* $\mathcal{E}$*. We call* $\mathcal{M}(\mathcal{E})$ *the* σ-**algebra generated by** $\mathcal{E}$ *. Moreover,* $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$ *implies that* $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(\mathcal{F})$.

Definition. *By the smallest* σ*-algebra above we mean* $\mathcal{M}(\mathcal{E})$ *is defined as the intersection of all* σ*-algebras containing* $\mathcal{E}$*. If* $\mathcal{M}$ *is a* σ*-algebra on* X*, we call the ordered pair* $(X, \mathcal{M})$ *a* **measurable space***, and the elements of* $\mathcal{M}$ *are often called the* **measurable sets** *of* X.

If (X, τ) is a topological space, then an important example is the space $(X, \mathcal{M}(\tau))$, where $\mathcal{M}(\tau)$ the σ-algebra generated by the topology τ. This σ-algebra is called the **Borel σ-algebra** on X and its members are called **Borel sets**.

We will later need the following notion of elementary family and how we can obtain an algebra from such families.

Definition. *A collection $\mathcal{E} \subset 2^X$ is called an* **elementary family** *if*

(i) $\emptyset \in \mathcal{E}$,

(ii) E, $F \in \mathcal{E}$ *implies* $E \cap F \in \mathcal{E}$,

(iii) $E \in \mathcal{E}$ *implies* E^c *is a finite disjoint union of members of* $\mathcal{E}$.

Proposition 1.1. *If $\mathcal{E}$ is an elementary family, then the collection $\mathcal{A}$ of finite disjoint unions of members of $\mathcal{E}$ is an algebra.*

1.1.2 Measurable functions

We let $(X, \mathcal{M})$ and $(X, \mathcal{N})$ are two measurable space, and suppose $f : X \longrightarrow Y$ is a mapping.

Definition. *The function $f : X \longrightarrow Y$ is called a* **measurable***, or more precisely $(\mathcal{M}, \mathcal{N})$-measurable, if $f^{-1}(E) \in \mathcal{M}$ for each $E \in \mathcal{N}$.*

Proposition 1.2. *If $(X, \mathcal{M})$, $(Y, \mathcal{N})$ and $(Z, \mathcal{O})$ are measure spaces, and if $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are measurable functions, then the composition mapping $h : g \circ f : X \longrightarrow Z$ is measurable.*

Proof. Indeed, if $E \in \mathcal{O}$, then $\tilde{E} := g^{-1}(E) \in \mathcal{N}$ since g is measurable. Thus, $h^{-1}(E) = f^{-1} \circ g^{-1}(E) = f^{-1}(g^{-1}(E)) = f^{-1}(\tilde{E}) \in \mathcal{M}$. □

Proposition 1.3. *If $\mathcal{N}$ is generated by $\mathcal{E}$, then $f : X \longrightarrow Y$ is $(\mathcal{M}, \mathcal{N})$-measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for each $E \in \mathcal{E}$.*

Proof. The forward implication is immediate. To prove the converse, notice that $\{E \subset Y \mid f^{-1}(E) \in \mathcal{M}\}$ is a σ-algebra containing $\mathcal{E}$. Therefore, $\mathcal{M}(\mathcal{E}) \subset \{E \subset Y \mid f^{-1}(E) \in \mathcal{M}\}$ and thus, if $E \in \mathcal{N}$, then $f^{-1}(E) \in \mathcal{M}$. □

The previous two propositions imply continuous functions between topological spaces are measurable.

Corollary 1.1. *If $(X, \mathcal{M}(\tau_X))$ and $(Y, \mathcal{M}(\tau_Y))$ are measurable spaces, and $f : X \longrightarrow Y$ is continuous on X, then f is measurable.*

1.1.3 Measures

Let $(X, \mathcal{M})$ be a measurable space.

Definition. *A* **measure***, or more precisely a* **positive measure***, on* $(X, \mathcal{M})$*, or simply on* X *if* $\mathcal{M}$ *is understood, is a function* $\mu : \mathcal{M} \longrightarrow [0, +\infty]$ *such that*

(a) $\mu(\emptyset) = 0$,

(b) if $\{E_j\}_{j=1}^{\infty}$ *is a sequence of disjoint sets in* $\mathcal{M}$*, then*

$$\mu(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{\infty} \mu(E_i), \tag{1.1}$$

where the last condition is commonly referred to as the **countable additivity** *property.*

Note that the infinite sum in (1.1) is allowed to take the value $+\infty$, e.g., if either $\mu(E_{i_0}) = +\infty$ for some $i_0 \in \mathbb{N}$, or the right-hand side yields a divergent sum of finite, non-negative numbers.

The following properties is a simple exercise we leave for the reader to verify.

Proposition 1.4. *Let* $\{E_i\}$ *be any sequence of measurable sets in* $\mathcal{M}$*. Then,*

(a) $E_1 \subset E_2$ *implies that* $\mu(E_1) \leq \mu(E_2)$;

(b) if $E_1 \subset E_2 \subset E_3 \subset \ldots$, *then* $\lim_{i\to\infty} \mu(E_i) = \mu(\cup_{i=1}^{\infty} E_i)$; *and*

(c) if $\mu(E_1) < +\infty$ *and* $E_1 \supset E_2 \supset E_3 \supset \ldots$, *then* $\lim_{i\to\infty} \mu(E_i) = \mu(\cap_{i=1}^{\infty} E_i)$.

Definition. *If* μ *is a measure on* $(X, \mathcal{M})$*, then the ordered triple* $(X, \mathcal{M}, \mu)$ *is called a* **measure space***.*

Definition. *Given a measure space* $(X, \mathcal{M}, \mu)$*, we say* μ *is* **finite** *if* $\mu(X) < +\infty$*, and we say a set* $E \subset X$ *is* σ**-finite** *with respect to* μ *provided that* $E = \cup_{i=1}^{\infty} E_i$ *for some measurable sets* $E_i \in \mathcal{M}$ *such that* $\mu(E_i) < +\infty$*, for each* $i \in \mathbb{N}$*. In the particular case* $E = X$*, we just say* μ *is* σ**-finite***.*

In the case X is a topological space and $\mathcal{M} = \mathcal{B}_X$ is the Borel σ-algebra on X, then

Definition. *A measure on* $(X, \mathcal{B}_X)$ *is called a* **Borel measure***.*

1.1.4 Null sets and complete measures

Again, suppose $(X, \mathcal{M}, \mu)$ is a given measure space. An important class of non-trivial measurable sets are those that have measure zero, which we call **null sets**. Obviously, countable subadditivity implies the countable union of null sets is null, a simple consequence we use often enough. It will also be common and necessary to assert a statement is true or a certain property holds for all elements x in X except in some null set. In such a situation, we say the statement or property is true **almost everywhere**, **μ-almost everywhere**, or we simply just write **a.e.** for short.

It is clear that part (a) of Proposition 1.4 ensures all subsets F of a null set $E \in \mathcal{M}$ is null so long as F is measurable, i.e., $F \in \mathcal{M}$. In developing a practical and robust theory, we cannot assume all sets in 2^X are measurable but we adopt the following notion of completeness for measures.

Definition. *A measure whose $\mu : \mathcal{M} \longrightarrow [0, +\infty]$ whose domain $\mathcal{M}$ includes all subsets of null sets is said to be* **complete***, or more precisely, the measure space $(X, \mathcal{M}, \mu)$ is said to be* **complete***.*

Completeness is a convenient property to have in practice, and fortunately, we can always extend any given measure space to one with a complete measure, if needed.

Theorem 1.1 (Completeness)**.** *Suppose $(X, \mathcal{M}, \mu)$ is a measure space and set*

$$\overline{\mathcal{M}} = \big\{ E \cup F \,|\, E \in \mathcal{M},\ F \subset N \text{ for some null set } N \in \mathcal{M} \big\}.$$

Then

(a) $\overline{\mathcal{M}}$ is a σ-algebra on X, and

(b) there exists a unique complete measure $\overline{\mu}$ on $\overline{\mathcal{M}}$ such that $\overline{\mu}\big|_{\mathcal{M}} = \mu$.

Remark. *The unique extension $\overline{\mu}$ in the previous theorem is called the* **completion** *of the μ, and $\overline{\mathcal{M}}$ is called the* **completion** *of $\mathcal{M}$ with respect to μ.*

1.2 Outer measures

In this section, we outline an approach for constructing measures from a weaker set of conditions and a mapping defined on the entire power set 2^X. Specifically, given a non-empty set X, an **outer measure** on X is a function $\mu^* : 2^X \longrightarrow [0, +\infty]$ such that

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

(b) if $A \subset B$, then $\mu^*(A) \leq \mu^*(B)$, (monotonicity)

(c) $\mu^*(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$ for any sequence $\{A_i\}_{i=1}^{\infty}$ in 2^X. (countable subadditivity)

The following result asserts that starting with any sub-collection in 2^X containing both the empty set and entire set, and given a reasonably defined function on it, then we can always generate an outer measure on X.

Theorem 1.2. *If we have a collection $\mathcal{E} \subset 2^X$ and a function $\rho : \mathcal{E} \longrightarrow [0, +\infty]$ such that $\emptyset, X \in \mathcal{E}$, and $\rho(\emptyset) = 0$, then the function $\mu^* : 2^X \longrightarrow [0, +\infty]$, where*

$$\mu^*(A) = \inf \Big\{ \sum_{i=1}^{\infty} \rho(E_i) \,\Big|\, E_i \in \mathcal{E},\, A \subset \cup_{i=1}^{\infty} E_i \Big\},$$

defines an outer measure on X.

The next result reveals how to start with an outer measure then construct a complete measure space. Some definitions and motivating comments are in order.

Definition. *If μ^* is an outer measure on a non-empty set X, a set $A \subset X$ is said to be* **μ^*-measurable** *if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \ \ \text{for all } E \subset X.$$

In fact, we may weaken and replace the equality in this definition with the inequality

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

since the reverse inequality automatically holds thanks to the countable subadditivity property of outer measures. Now, we come to the claimed result.

Theorem 1.3 (Carathèodory)**.** *If μ^* is an outer measure on X, the collection $\mathcal{M}$ of μ^*-measurable sets forms a σ-algebra on X, and the restriction of μ^* to $\mathcal{M}$ defines a complete measure.*

An application of Carathèdory's theorem will lead to a complete measure space provided we start with a "premeasure" on only an algebra on X.

Definition. *If $\mathcal{A} \subset 2^X$ is an algebra on X, a function $\mu_0 : \mathcal{A} \longrightarrow [0, +\infty]$ is called a* **premeasure** *on $\mathcal{A}$ if*

(a) $\mu_0(\emptyset) = 0$,

(b) if $\{A_i\}_{i=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$, then

$$\mu_0\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu_0(A_i).$$

A premeasure μ_0 on $\mathcal{A}$ induces an outer measure μ^* on X ensured by Theorem 1.2, i.e.,

$$\mu^*(A) = \inf\Big\{ \sum_{i=1}^{\infty} \mu_0(E_i) \,\Big|\, E_i \in \mathcal{E},\ A \subset \cup_{i=1}^{\infty} E_i \Big\}. \tag{1.2}$$

Proposition 1.5. *If μ_0 is a premeasure on an algebra $\mathcal{A}$ and μ^* is defined in* (1.2)*, then*

(i) $\mu^*\big|_{\mathcal{A}} = \mu_0$*, and*

(ii) every set in $\mathcal{A}$ is μ^*-measurable.

Theorem 1.4. *Let $\mathcal{A} \subset 2^X$ be an algebra on X, μ_0 is a premeasure on $\mathcal{A}$, and set $\mathcal{M} := \mathcal{M}(\mathcal{A})$, the σ-algebra generated by $\mathcal{A}$. Then,*

(a) There exists a measure μ on $\mathcal{M}$ such that $\mu = \mu^*\big|_{\mathcal{M}}$, where μ^* is defined in (1.2)*.*

(b) If ν is another measure on $\mathcal{M}$ that extends μ_0, i.e., $\nu\big|_{\mathcal{A}} = \mu_0$, then

$$\nu(E) \le \mu(E) \ \text{ for all } E \in \mathcal{M} \ \text{ with equality if } \mu(E) < \infty.$$

(c) If μ_0 is σ-finite in the same sense with measures, then μ is the unique extension of μ_0 to a measure on $\mathcal{M}$.

1.3 Monotone classes and the uniqueness of measures

Definition. *A* **monotone class** *$\mathcal{M}$ is a collection of sets satisfying the following two properties.*

(a) If $\{A_i\}_{i=1}^n$ is a sequence of sets in $\mathcal{M}$ and $A_1 \subset A_2 \subset A_3 \subset \ldots$, then $\bigcup_{i=1}^\infty A_i \in \mathcal{M}$.

(b) If $\{B_i\}_{i=1}^n$ is a sequence of sets in $\mathcal{M}$ and $B_1 \supset B_2 \supset B_3 \supset \ldots$, then $\bigcap_{i=1}^\infty B_i \in \mathcal{M}$.

The following result asserts σ-algebras are monotone limits of algebras.

Theorem 1.5 (Monotone Class)**.** *Let X be a non-empty set and let $\mathcal{E}$ be an algebra of sets on X such that $\emptyset, X \in \mathcal{E}$. Then, there exists a smallest monotone class $\mathcal{M}$ that contains $\mathcal{E}$. Furthermore, $\mathcal{M}$ is also the smallest σ-algebra containing $\mathcal{E}$, i.e., $\mathcal{M} = \mathcal{M}(\mathcal{E})$.*

One of many important applications of the Monotone Class Theorem is the uniqueness of measures.

Theorem 1.6. *Let X be a non-empty set and $\mathcal{E}$ is an algebra of sets on X, and assume $\emptyset, X \in \mathcal{E}$, and recall that $\mathcal{M}(\mathcal{E})$ denotes the σ-algebra generated by $\mathcal{E}$. Let μ be a σ-finite measure in the stronger sense that there is a sequence of sets $\{E_i\}_{i=1}^\infty$ in $\mathcal{E} \subset \mathcal{M}(\mathcal{E})$, with $\mu(E_i) < +\infty$ for each $i \in \mathbb{N}$, such that $X = \bigcup_{i=1}^\infty E_i$. If ν is a measure that coincides with μ on $\mathcal{E}$, then $\mu = \nu$ on all of $\mathcal{M}(\mathcal{E})$.*

Proof. □

1.4 A model case: Borel measures on the real line

We focus on the special case $X = \mathbb{R}$ with Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ generated by the open sets of $\mathbb{R}$. We characterize the Borel sets in $\mathbb{R}$ and study the measures on these sets. In fact, the Borel sets of $\mathbb{R}$ can be generated by open intervals, closed intervals, half-open intervals, etc. For instance, there holds

Proposition 1.6. *If $\mathcal{E} = \{(a,\infty) \mid a \in \mathbb{R}\}$, then $\mathcal{M}(\mathcal{E}) = \mathcal{B}_{\mathbb{R}}$.*

Proof. Indeed, each member of $\mathcal{E}$ is an open set in $\mathbb{R}$ and so $\mathcal{E} \subset \mathcal{B}_{\mathbb{R}}$, which further implies $\mathcal{M}(\mathcal{E}) \subset \mathcal{B}_{\mathbb{R}}$ by Lemma 1.1. Conversely, since the family of open intervals forms a base for the topology $\tau_{\mathbb{R}}$ of $\mathbb{R}$, and since all open intervals are contained in $\mathcal{M}(\mathcal{E})$ by construction, Lemma 1.1 also implies $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\tau_{\mathbb{R}}) \subset \mathcal{M}(\mathcal{E})$. □

Remark. *As hinted earlier, we can replace $\mathcal{E}$ with other families of unbounded intervals such as*

$$\{(a,\infty) \mid a \in \mathbb{R}\},\ \{[a,\infty) \mid a \in \mathbb{R}\},\ \{(-\infty,b) \mid b \in \mathbb{R}\} \quad or \quad \{(-\infty,b] \mid b \in \mathbb{R}\};$$

or even families consisting of bounded intervals such as

$$\{(a,b) \mid a < b \text{ in } \mathbb{R}\},\ \{[a,b] \mid a < b \text{ in } \mathbb{R}\},\ \{[a,b) \mid a < b \text{ in } \mathbb{R}\} \quad or \quad \{(a,b] \mid a < b \text{ in } \mathbb{R}\}.$$

Verifying these assertions, however, is left to the reader.

To motivate what follows, suppose we have a finite Borel measure μ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and we define a mapping $F : \mathbb{R} \longrightarrow [0,\infty)$, where $F(x) = \mu((-\infty, x])$. Then F is monotone increasing and right continuous in $\mathbb{R}$, and for $a < b$,

$$\mu((a,b]) = F(b) - F(a). \tag{1.3}$$

Notice the case when $F(x) = x$ is the measure that assigns the length of the interval. Interestingly, we will show the reverse process holds, i.e., any monotone increasing and right continuous function induces a unique Borel measure satisfying "nice" properties such as (1.3). First, we shall require the following result whose proof, although tedious, is straightforward and left for the reader (c.f. page 33 in Folland). Now, we consider the collection of intervals of the form $\emptyset$, $(a,b]$, or (a,∞), where $-\infty \le a < b < \infty$. We call such intervals ***h*-intervals**, and we note this collection is closed under disjoint unions, complements, and intersections. We define $\mathcal{A}$ to be the collection of finite disjoint unioins of h-intervals, which defines an algebra on $\mathbb{R}$ by Proposition 1.1. Proposition 1.1 further implies that $\mathcal{M}(\mathcal{A}) = \mathcal{B}_{\mathbb{R}}$.

Proposition 1.7. *Let $\mathcal{A}$ be the algebra of finite disjoint unions of h-intervals, and suppose that $F : \mathbb{R} \longrightarrow \mathbb{R}$ is monotone increasing and right continuous. If $\{(a_i, b_i] \mid i = 1, 2, \ldots, n\}$ are disjoint h-intervals, we define $\mu_0(\emptyset) = 0$ and*

$$\mu_0\big(\cup_{i=1}^n (a_i, b_i]\big) = \sum_{i=1}^n F(b_i) - F(a_i).$$

Then μ_0 is a premeasure on $\mathcal{A}$.

Theorem 1.7. *If $F : \mathbb{R} \longrightarrow \mathbb{R}$ is any monotone increasing and right continuous function, then there exists a unique Borel measure μ_F on $\mathbb{R}$ such that*

$$\mu_F((a,b]) = F(b) - F(a) \quad \textit{for all} \quad a, b \in \mathbb{R}.$$

If G is another such function, then $\mu_F = \mu_G$ if and only if $F - G$ is constant in $\mathbb{R}$.

Conversely, if μ is a Borel measure on $\mathbb{R}$ that is finite on all Borel sets, and we define $F : \mathbb{R} \longrightarrow \mathbb{R}$ to satisfy $F(0) = 0$, $F(x) = \mu((0,x])$ if $x > 0$, and $F(x) = -\mu((x,0])$ if $x < 0$, then F is monotone increasing and right continuous, and μ coincides with μ_F.

CHAPTER 2

Abstract Integration

2.1 Integration of non-negative functions

Suppose we are given a measure space $(X, \mathcal{M}, \mu)$, and we would like to define a notion of an integral of a given function defined on our space. We shall construct such an integral in this section but before we can do so, we must clarify the class of functions suitable in our construction. These functions are the so-called measurable functions from the previous chapter.

If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are a pair of measurable spaces, recall we defined the notion of a measurable mapping between them as follows. A function $f : X \longrightarrow Y$ is called $(\mathcal{M}, \mathcal{N})$-measurable or just measurable if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$. In the special case when X and Y are topological spaces whose σ-algebras are generated by their topologies, then the continuous functions mapping X into Y are measurable thanks to Corollary 1.1. Another special case is if $Y = \mathbb{R}$, in which case a real-valued mapping $f : X \longrightarrow \mathbb{R}$ is evidently $\mathcal{M}$-measurable if and only if $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. In fact, thanks to Proposition 1.6, we can replace the interval (a, ∞) in this characterization with any one of the following other types: $[a, \infty)$, $(-\infty, a)$ or $(-\infty, a]$.

Here, we shall mostly focus on extended real-valued measurable functions on X in defining our abstract notion of integration. A special class of measurable functions are the simple functions, which we will use to construct our integral over a measure space X. By a **simple function** on X, we mean a finite linear combination of characteristic functions of sets in the σ-algebra $\mathcal{M}$ (where we do not allow such simple functions to assume values of $\pm\infty$). In other words, a simple function φ is measurable and can be represented in the form

$$\varphi = \sum_{i=1}^{n} a_i \chi_{E_i},$$

where $a_i \in \mathbb{R}$ and χ_{E_i} is the characteristic function of a measurable set $E_i \in \mathcal{M}$. Indeed, this representation of φ is not necessarily unique, however there is a unique **standard representation** characterized by the fact that the coefficients a_i are distinct and the measurable sets E_i are disjoint non-empty subsets such that $X = \cup_{i=1}^n E_i$.

We are now prepared to define the integral of a measurable function by first defining the integral of a simple function.

Definition. *Given a measure space $(X, \mathcal{M}, \mu)$, L^+ will denote the space of all $\mathcal{M}$-measurable functions from X into $[0, +\infty]$.*

Definition (Integral of a simple function)**.** *If φ is a simple function in L^+ with standard representation $\varphi = \sum_{i=1}^n a_i \chi_{E_i}$, then the* **integral** *of φ with respect to the measure μ is defined by*

$$\int_X \varphi \, d\mu = \sum_{i=1}^n a_i \mu(E_i),$$

(with the convention that $0 \cdot +\infty = 0$ on the right-hand side).

With this definition, the integral of any non-negative measurable function is obtained as an approximation from below by simple functions.

Definition. *In general, if $f \in L^+$, the* **integral** *of f with respect to μ is defined by*

$$\int_X f \, d\mu = \sup \left\{ \int_X \varphi \, d\mu \,\middle|\, \varphi \text{ is simple, } 0 \leq \varphi \leq f \right\}.$$

Remark. *For brevity, we sometimes use the short-hand notation $\int f \, d\mu$ or $\int f$ in place of $\int_X f \, d\mu$ when the space X and measure μ are obvious. Some authors also use the conventional notation $\int_X f(x) \, d\mu(x)$ or $\int_X f(x) \, \mu(dx)$ to denote the integral.*

Definition. *If $E \in \mathcal{M}$, then $f\chi_E$ is also in L^+ and we define the* **integral** *of f on E by*

$$\int_E f \, d\mu = \int_X f \chi_E \, d\mu.$$

The following familiar results (for the Riemann integral) hold for simple functions and can easily be extended to functions in L^+. The details are left to the reader.

Proposition 2.1. *Let φ and ψ be simple functions in L^+. Then the following statements hold.*

(a) If $c \geq 0$, then $\int (c\varphi) \, d\mu = c \int \varphi \, d\mu$;

(b) $\displaystyle\int (\varphi + \psi)\, d\mu = \int \varphi\, d\mu + \int \psi\, d\mu$;

(c) $\varphi \leq \psi$ *implies* $\displaystyle\int \varphi\, d\mu \leq \int \psi\, d\mu$;

(d) the mapping $E \mapsto \displaystyle\int_E \varphi\, d\mu$ *defines a measure on* $\mathcal{M}$.

There is an important class of measurable functions that will appear regularly hereafter, and these are the Lebesgue integrable functions. More generally, we consider complex measurable functions $f : X \longrightarrow \mathbb{C}$ on a given measure space $(X, \mathcal{M}, \mu)$. Indeed $|f|$ is a measurable function from X into $[0, +\infty]$ and so its integral $\int_X |f|\, d\mu$ is defined. Then, we denote by $L^1(\mu)$ the set of all complex measurable functions on X for which

$$\int_X |f|\, d\mu < +\infty,$$

and we call the elements of $L^1(\mu)$ the **summable** or the **Lebesgue integrable** functions (with respect to μ). For $1 \leq p < +\infty$ we define by $L^p(\mu)$ the set of all complex measurable functions $f : X \longrightarrow \mathbb{C}$ such that

$$\int_X |f|^p\, d\mu < +\infty,$$

and the elements of $L^p(\mu)$ are called the p-**summable** functions (with respect to μ). The set $L^p(\mu)$ is indeed a linear space in which

$$\|f\|_{L^p(\mu)} := \Big(\int_X |f|^p\, d\mu \Big)^{1/p}$$

defines a norm on it. In fact, $L^p(\mu)$ defines a Banach space under this norm, and $L^p(\mu)$ is commonly called a **Lebesgue** or L^p space.

2.2 The convergence theorems

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

Theorem 2.1 (Lebesgue's Montone Convergence)**.** *Let* $\{f_n\}$ *be a* **monotone increasing** *sequence of non-negative measurable functions that converges* **pointwise** *to a function* $f(x)$*, i.e.,*

(a) $0 \leq f_1(x) \leq f_2(x) \leq \ldots \leq f_n(x) \leq \ldots \leq \infty$ *for every* $x \in X$ *(monotone increasing),*

(b) and

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for every } x \in X \text{ (pointwise convergence).}$$

Then f is measurable and

$$\int_X f_n \, d\mu \longrightarrow \int_X f \, d\mu \quad \text{as } n \longrightarrow \infty. \tag{2.1}$$

Proof. Since the pointwise limit of a sequence of non-negative measurable functions is also a non-negative measurable function, the limiting function $f : X \longrightarrow [0, \infty]$ is also measurable. Moreover, since $f_n \leq f_{n+1} \leq f$ for all $n \in \mathbb{N}$, we deduce that

$$\int_X f_n \, d\mu \leq \int_X f_{n+1} \, d\mu \leq \int_X f \, d\mu \quad \text{for all } n \in \mathbb{N}.$$

Thus,

$$\lim_{n \to \infty} \int_X f_n \, d\mu \leq \int_X f \, d\mu. \tag{2.2}$$

To obtain the reverse inequality, we choose an arbitrary $0 < \alpha < 1$ and let φ be any simple function satisfying $0 \leq \varphi \leq f$. Set

$$A_n := \Big\{ x \in X \mid f_n(x) \geq \alpha \varphi(x) \Big\}.$$

It is easy to see that $A_n \in \mathcal{M}$ and $A_n \subset A_{n+1}$ for each n, and that $X = \cup_{n=1}^{\infty} A_n$. Under these conditions on $\{A_n\}_{n=1}^{\infty}$, Proposition 1.4 implies that

$$\nu(\cup_{n=1}^{\infty} A_n) = \lim_{n \to \infty} \nu(A_n)$$

for any measure ν on $(X, \mathcal{M})$, and it is also standard to show $\nu(E) := \int_E \varphi \, d\mu$ indeed provides such a particular measure (actually, we will soon show in Corollary 2.2 that this remains valid if φ is any non-negative measurable function and not just a simple function). These standard results imply that

$$\int_X \varphi \, d\mu = \nu(X) = \nu(\cup_{n=1}^{\infty} A_n) = \lim_{n \to \infty} \nu(A_n) = \lim_{n \to \infty} \int_{A_n} \varphi \, d\mu.$$

On the other hand, by our construction of $\{A_n\}_{n=1}^{\infty}$, we have that

$$\alpha \int_{A_n} \varphi \, d\mu = \int_{A_n} \alpha \varphi \, d\mu \leq \int_{A_n} f_n \, d\mu \leq \int_X f_n \, d\mu.$$

Sending $n \longrightarrow \infty$ in the previous inequality leads us to

$$\alpha \int_X \varphi \, d\mu = \lim_{n \to \infty} \int_{A_n} \alpha \varphi \, dx \leq \lim_{n \to \infty} \int_X f_n \, d\mu. \tag{2.3}$$

Recall that the integral of f is defined by

$$\int_X f \, d\mu = \sup \int_X \phi \, d\mu,$$

where the supremum is taken over all simple functions ϕ such that $0 \leq \phi \leq f$. Therefore, since $0 < \alpha < 1$ and φ were chosen arbitrarily and because of (2.3), we must have that

$$\int_X f \, d\mu \leq \lim_{n\to\infty} \int_X f_n \, d\mu.$$

Combining this with (2.2) completes the proof of the theorem. □

Remark. *The integral in* (2.1) *is allowed to equal* $+\infty$. *Moreover, an analogue result involving non-increasing sequences of functions holds true. Namely, if there is a sequence* $\{f_n\}_{n=1}^{\infty}$ *of non-negative measurable functions with* $f_1 \in L^1(\mu)$ *and this sequence is non-decreasing, i.e.,* $f_1(x) \geq f_2(x) \geq \ldots \geq f(x)$ *for all* $x \in X$, *and if* $f_n(x) \longrightarrow f(x)$ *for all* $x \in X$, *then* $f : X \longrightarrow [0, \infty]$ *is measurable and*

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

The assumption that $f \in L^1(\mu)$ *cannot be omitted.*

We recall several important applications and consequences of the Monotone Convergence Theorem. The first is Fatou's lemma.

Lemma 2.1 (Fatou's)**.** *If* $f_n : X \longrightarrow [0, \infty]$ *is measurable for each positive integer* n, *then*

$$\int_X \left(\liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Proof. Set $g_n = \inf_{k\geq n} f_k$ and set $g := \liminf_{n\to\infty} f_n$. Then $g_n : X \longrightarrow [0, \infty]$ is measurable, $g_n \longrightarrow g$ pointwise everywhere in X and $\{g_n\}$ is monotone increasing. By the Monotone Convergence Theorem and the fact that $f_n \geq g_n$ in X, for all n, we obtain

$$\int_X \left(\liminf_{n\to\infty} f_n \right) d\mu = \int_X g \, d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

□

The next is a consequence of Fatou's lemma which we often use. For instance, it implies that strong solutions of elliptic equations on a bounded domain satisfy the equation pointwise almost everywhere in the domain.

Corollary 2.1. *Suppose that* f *is a non-negative measurable function. Then* $f = 0$ μ*-almost everywhere in* X *if and only if*

$$\int_X f \, d\mu = 0. \tag{2.4}$$

Proof. If (2.4) holds, let

$$E_n = \left\{ x \in X \,\middle|\, f(x) > 1/n \right\},$$

so that E_n is measurable and $f \geq (1/n)\chi_{E_n}$, from which we see

$$0 = \int_X f\, d\mu \geq \frac{1}{n}\mu(E_n) \geq 0.$$

Thus, $\mu(E_n) = 0$ and so the set

$$\{x \in X \mid f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$$

is measurable and has measure zero by the countable additive property of measures. This verifies that $f = 0$ μ-almost everywhere in X.

Conversely, assume $f = 0$ μ-almost everywhere. If

$$E = \{x \in X \mid f(x) > 0\},$$

then obviously E is measurable with $\mu(E) = 0$. Then set $f_n = n\chi_E$ so that each f_n is non-negative and measurable, and clearly $f \leq \liminf_{n\to\infty} f_n$. Thus, by Fatou's lemma,

$$0 \leq \int_X f\, d\mu \leq \liminf_{n\to\infty} \int_X f_n\, d\mu = \liminf_{n\to\infty} n\mu(E) = 0.$$

Hence, $\|f\|_{L^1(\mu)} = 0$, and this completes the proof. □

The next consequence illustrates we can use the integral of any non-negative measurable function to construct another measure that is absolutely continuous with respect to the original measure.

Corollary 2.2. *If $f : X \longrightarrow [0, \infty]$ is a non-negative measurable function and if λ is defined on the σ-algebra $\mathcal{M}$ by*

$$\lambda(E) = \int_E f\, d\mu, \tag{2.5}$$

then λ is a measure on the measurable space $(X, \mathcal{M})$. Moreover, the measure λ is absolutely continuous with respect to μ in the sense that if $E \in \mathcal{M}$ and $\mu(E) = 0$, then $\lambda(E) = 0$.

Proof. We verify λ defines a measure. Obviously, $\lambda(\emptyset) = 0$. Now, suppose that $\{E_n\}_{n=1}^{\infty} \subset \mathcal{M}$ is a sequence of disjoint measurable sets. Set $E := \cup_{n=1}^{\infty} E_n \in \mathcal{M}$ and define

$$f_n(x) = \sum_{k=1}^{n} f\chi_{E_k}.$$

Indeed, f_n is a non-negative measurable function and

$$\int_X f_n \, du = \sum_{k=1}^{n} \int_X f\chi_{E_k} \, d\mu = \sum_{k=1}^{n} \lambda(E_k).$$

Then $\{f_n\}$ is a monotone increasing sequence of non-negative, measurable functions converging pointwise to f on X. Hence, the Monotone Convergence Theorem implies that

$$\lambda(E) = \int_E f \, du = \lim_{n\to\infty} \int_X f_n \, d\mu = \sum_{n=1}^{\infty} \lambda(E_n),$$

and therefore λ defines a measure.

Assume now that $E \in \mathcal{M}$ such that $\mu(E) = 0$. The function $f\chi_E$ vanishes μ-almost everywhere. So, by Corollary 2.1, we deduce that

$$\lambda(E) = \int_X f\chi_E \, d\mu = 0.$$

□

Remark. *In general, we write $\lambda \ll \mu$ to mean λ is absolutely continuous with respect to μ. Under suitable conditions, the converse of Corollary 2.2 holds, and this result is well-known and is referred to as the Radon-Nikodym Theorem. We state it below for completeness but omit its proof. The proof can be found in any standard graduate real analysis textbook, e.g., see our references [2, 3, 4].*

Theorem 2.2 (Radon-Nikodym)**.** *Let λ and μ be σ-finite measures on $(X, \mathcal{M})$ and suppose $\lambda \ll \mu$. Then there exists a measuable function $f : X \longrightarrow [0, \infty]$ such that*

$$\lambda(E) = \int_E f \, d\mu, \; E \in \mathcal{M}.$$

Moreover, the function f is uniquely determined μ-almost everywhere.

Remark. *The function f in Theorem 2.2 is called the Radon-Nikodym derivative of λ with respect to μ and we write*

$$\frac{d\lambda}{d\mu} = f.$$

We can invoke the earlier corollary to replace pointwise convergence with μ-almost everywhere convergence in Theorem 2.1 but the limit function is assumed to be measurable a priori.

Corollary 2.3. *Let $\{f_n\}$ be a monotone increasing sequence of non-negative measurable functions that converges μ-almost everywhere in X to a non-negative measurable function $f(x)$. Then*

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

Proof. Choose $N \in \mathcal{M}$ be such that $\mu(N) = 0$ and $\{f_n\}$ converges to f at every point of $M = X \backslash N$. Then $\{f_n \chi_M\}$ converges to $f \chi_M$ in X. Thus Theorem 2.1 implies that

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

Since $\mu(N) = 0$, the functions $f \chi_N$ and $f_n \chi_N$ vanish μ-almost everywhere. It follows from 2.4 that

$$\int_X f \chi_N \, d\mu = 0 \quad \text{and} \quad \int_X f_n \chi_N \, d\mu = 0.$$

Since $f = f \chi_M + f \chi_N$ and $f_n = f_n \chi_M + f_n \chi_N$, it follows that

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

□

An essential convergence theorem often utilized in our applications is Lebesgue's Dominated Convergence Theorem (LDCT). This useful result simply follows from Fatou's lemma and the following basic fact.

Lemma 2.2. *If $f \in L^1(\mu)$, then*

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

Proof. Set $z = \int_X f \, d\mu \in \mathbb{C}$. Thus, $|z| = \alpha z$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. If $u = Re(\alpha f)$, then $u \leq |\alpha f| = |f|$ and so

$$|z| = \alpha z = \alpha \int_X f \, d\mu = \int_X \alpha f \, d\mu = \int_X u \, d\mu \leq \int_X |f| \, d\mu,$$

where we used the fact that $\int_X \alpha f \, d\mu$ is real. □

Theorem 2.3 (Lebesgue's Dominated Convergence)**.** *Suppose $\{f_n\}$ is a sequence of measurable functions on X such that*

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

exists for every $x \in X$. If there is a function $g \in L^1(\mu)$ such that

$$|f_n(x)| \leq g(x) \quad \text{for } n = 1, 2, 3, \ldots; \, x \in X,$$

then $f \in L^1(\mu)$,

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

and

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

Proof. Indeed, $f \in L^1(\mu)$, since $|f| \le g \in L^1(\mu)$ and f is measurable. We similarly deduce that $f_n \in L^1(\mu)$ for all n.

From the triangle inequality, we also get that $|f_n - f| \le 2g$, and so $f_n - f \in L^1(\mu)$. Applying Fatou's lemma to the non-negative functions $2g - |f_n - f|$ leads us to

$$\begin{aligned}\int_X 2g\,d\mu &\le \liminf_{n\to\infty} \int_X (2g - |f_n - f|)\,d\mu \\ &= \int_X 2g\,d\mu + \liminf_{n\to\infty} \Big(- \int_X |f_n - f|\,d\mu \Big) \\ &= \int_X 2g\,d\mu - \limsup_{n\to\infty} \int_X |f_n - f|\,d\mu.\end{aligned}$$

Noting that $\int_X g\,d\mu$ is finite, we may add $-2\int_X g\,d\mu$ to previous inequality to arrive at

$$\limsup_{n\to\infty} \int_X |f_n - f|\,d\mu \le 0.$$

This further implies that

$$\lim_{n\to\infty} \int_X |f_n - f|\,d\mu = 0. \tag{2.6}$$

Since $f_n - f \in L^1(\mu)$, Lemma 2.2 implies that

$$\Big| \int_X (f_n - f)\,d\mu \Big| \le \int_X |f_n - f|\,d\mu,$$

and so (2.6) further yields that

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

□

Remark. *In Theorem 2.3, we can easily weaken the statement and only assume that pointwise convergence holds in the μ-almost everywhere sense. This is because we can always redefine f_n and f on a set of measure zero.*

More precisely, since a countable union of measurable sets of measure zero is measurable and also has measure zero, we can find a measurable set E with $\mu(E) = 0$ and redefine $\{f_n\}$, and similarly with f, so that $f_n(x) = 0$ for $x \in E$ and $f_n(x)$ remains unchanged for $x \notin E$. Note this does not change the value of the integrals $\int_X f_n\,d\mu$.

An immediate application of Theorem 2.3 is the following

Corollary 2.4. *If $t \longrightarrow f(x,t)$ is continuous on $[a,b]$ for each $x \in X$, and if there exists $g \in L^1(\mu)$ such that $|f(x,t)| \le g(x)$ for $x \in X$, then the function F defined by*

$$F(t) = \int_X f(x,t)\,d\mu(x) \tag{2.7}$$

is continuous for each t in $[a,b]$.

Another basic application of Theorem 2.3 indicates when we may differentiate F and when it is equivalent to passing derivatives onto the integrand f. Hereafter, an integrable function f on X means f is a measurable function on X belonging to $L^1(\mu)$.

Corollary 2.5. *Suppose that for some t_0 in $[a, b]$, the function $x \longrightarrow f(x, t_0)$ is integrable on X, that $\partial f/\partial t$ exists on $X \times [a, b]$, and that there exists an integrable function g on X such that*

$$\left|\frac{\partial f}{\partial t}(x, t)\right| \leq g(x).$$

Then the function F as defined in (2.7) is differentiable on $[a, b]$ and

$$\frac{dF}{dt}(t) = \frac{d}{dt}\int_X f(x, t)\, d\mu(x) = \int_X \frac{\partial f}{\partial t}(x, t)\, d\mu(x).$$

Proof. Let t be any point of $[a, b]$. If $\{t_n\}$ is a sequence in $[a, b]$ converging to t with $t_n \neq t$, then

$$\frac{\partial f}{\partial t}(x, t) = \lim_{n\to\infty} \frac{f(x, t_n) - f(x, t)}{t_n - t}, \; x \in X.$$

Therefore, the function $x \longrightarrow (\partial f/\partial t)(x, t)$ is measurable.

If $x \in X$ and $t \in [a, b]$, by the mean-value theorem, there exists s_1 between t_0 and t such that

$$f(x, t) - f(x, t_0) = (t - t_0)\frac{\partial f}{\partial t}(x, s_1).$$

Therefore,

$$|f(x, t)| \leq |f(x, t_0)| + |t - t_0| g(x),$$

which implies that the function $x \longrightarrow f(x, t)$ is integrable for each t in $[a, b]$. Hence, if $t_n \neq t$, then

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_X \frac{f(x, t_n) - f(x, t)}{t_n - t}\, d\mu(x).$$

Since this integrand is dominated by $g(x)$, we may apply Theorem 2.3 to conclude the desired result.

□

We can use Theorem 2.3 to establish a similar convergence result in the Lebesgue spaces $L^p(\mu)$ with $1 \leq p < \infty$.

Theorem 2.4. *Let $1 \leq p < \infty$ and suppose $\{f_n\}$ is a sequence in $L^p(\mu)$ which converges μ-almost everywhere to a measurable function f. If there exists a $g \in L^p(\mu)$ such that*

$$|f_n(x)| \leq g(x), \; x \in X, \; n \in N,$$

then f belongs to $L^p(\mu)$ and $\{f_n\}$ converges in $L^p(\mu)$ to f.

Proof. Assume $1 < p < \infty$ since the case $p = 1$ is exactly Theorem 2.3. Obviously, the following two properties hold for μ-almost everywhere,

$$|f_n(x) - f(x)|^p \leq [2g(x)]^p, \text{ and } \lim_{n\to\infty} |f_n(x) - f(x)|^p = 0;$$

and there holds $[2g]^p$ and thus g^p belongs to $L^1(\mu)$. Hence, from Theorem 2.3, we get

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

and this completes the proof of the theorem.

□

Remark. *Lebesgue's dominated convergence theorem and its extension provide sufficient conditions that guarantee when pointwise convergence of a sequence of measurable functions implies strong convergence in the L^p norm topology; namely, if the sequence of functions can be compared to an L^p function, then pointwise convergence implies L^p convergence. Conversely, L^p convergence does not generally imply pointwise convergence. We give an example below illustrating this.*

Let $X = [0, 1]$, the sigma algebra $\mathcal{M}$ are the Borel sets, and μ is the Lebesgue measure. Consider the ordered list of intervals

$[0, 1], [0, \frac{1}{2}], [\frac{1}{2}, 1], [0, \frac{1}{3}], [\frac{1}{3}, \frac{2}{3}], [\frac{2}{3}, 1], [0, \frac{1}{4}], [\frac{1}{4}, \frac{1}{2}], [\frac{1}{2}, \frac{3}{4}], [\frac{3}{4}, 1], [0, \frac{1}{5}], [\frac{1}{5}, \frac{2}{5}], \ldots$; let f_n be the characteristic function of the n^{th} interval on this list, and let f be identically zero. If $n > m(m+1)/2 = 1+2+\ldots+m$, then f_n is a characteristic function of an interval I whose measure is at most $1/m$. Hence,

$$\|f_n - f\|^p_{L^p(\mu)} = \int_X |f_n - f|^p \, d\mu = \int_X |f_n|^p \, d\mu = \int_X f_n \, d\mu = \mu(I) \leq 1/m,$$

and this shows $\{f_n\}$ converges in L^p to $f \equiv 0$.

On the other hand, if x is any point of $[0, 1]$, then the sequence of numbers $\{f_n(x)\}$ has a subsequence consisting only of 1's and another subsequence consisting of 0's. Therefore, the sequence $\{f_n\}$ does not converge at any point of $[0, 1]$! (although we may select a particular subsequence of $\{f_n\}$ which does converge to f).

The next result swaps the domination condition in the LDCT with finite measure and uniform integrability.

Theorem 2.5 (Vitali's Convergence Theorem)**.** *Let $(X, \mathcal{M}, \mu)$ be of finite measure, i.e., $\mu(X) < +\infty$, and suppose the sequence $\{f_n\}$ is uniformly integrable over X, i.e., for every $\epsilon > 0$, there exists $\delta > 0$ such that for each n, E measurable and $\mu(E) < \delta$ implies $\int_E |f_n| \, d\mu < \epsilon$. If $\{f_n\}$ converges pointwise μ-a.e. in X to f, then*

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

CHAPTER 3

Radon Measures

We let X be a locally compact Hausdorff (LCH) topological space, and recall $\mathcal{B}_X$ denotes the Borel sets of X. The objective of this chapter is to give an introduction to Radon measures and their properties, and important characterizations useful in applications.

3.1 Regular Borel and Radon measures

We start by introducing regularity conditions on Borel measures, which will be useful in approximating the measure of each Borel set.

Definition. *A Borel measure μ on X is said to be* **outer regular** *on a set $E \subset X$ if*

$$\mu(E) = \inf \left\{ \mu(U) \,\middle|\, U \text{ is open}, \ E \subset U \right\},$$

and it said to be **inner regular** *on E if*

$$\mu(E) = \sup \left\{ \mu(K) \,\middle|\, K \text{ is compact}, \ K \subset E \right\}.$$

A Borel measure is said to be **regular** *if it is both outer and inner regular on all Borel sets.*

Thus, regularity of a Borel measure roughly states that the measure of a Borel set can be approximated inside and outside, respectively, by compact and open sets. We introduce a special type of Borel measure which are locally finite on compact subsets and that are regular in a slightly weaker sense than that of Borel regularity.

Definition. *A* **Radon measure** *on X is a Borel measure that is finite on all compact sets of X, outer regular on all Borel sets, and inner regular on all open sets of X.*

Remark. *Note that we are restricting the inner regularity of Radon measures to only the open sets rather than to all Borel sets. However, as we shall see below (see Theorem 3.1), Radon measures are inner regular on all Borel sets if we further assume X is σ-compact. If σ-compactness is dropped, then inner regularity on Borel sets may no longer hold (see the remark immediately after the corollary).*

3.2 The Riesz Representation Theorem

There is a profoundly deep connection between Radon measures and continuous functions with compact support, which is the topic of our next main result. Before we state the result, we first introduce some necessary notation and definitions. We let $C_c(X)$ denote the normed linear space of continuous (real-valued) functions on X with compact support. A norm on $C_c(X)$ is given by the uniform norm,

$$\|f\|_u := \sup_X |f(x)|.$$

The **dual space** of $C_c(X)$, denoted by $C_c(X)^*$, is the collection of all continuous (i.e. bounded) linear functionals on X. A **linear functional** $I : C_c(X) \longrightarrow \mathbb{R}$ is said to be **positive** if $f \geq 0$ implies that $I(f) \geq 0$. If U is an open set in X and $f \in C_c(X)$, we write

$$f \prec U$$

to mean that $0 \leq f \leq 1$ and *supp*$(f) \subset U$. If K is a compact set in X and $f \in C_c(K)$, we write

$$K \prec f$$

to mean $0 \leq f \leq 1$ in X, and $f \equiv 1$ in K.

The following is a basic boundedness property for positive linear functionals on X.

Proposition 3.1. *If $I : C_c(X) \longrightarrow \mathbb{R}$ is a positive linear functional, then for each compact subset K of X, there exists a constant $C(K)$, depending on K, such that $|I(f)| \leq C(K)\|f\|_u$ for all $f \in C_c(X)$ such that supp$(f) \subset K$.*

To motivate the following, we suppose μ is a Borel measure on X that is locally finite, i.e., $\mu(K) < +\infty$ for every compact set $K \subset X$. Then, it is obvious $C_c(X)$ is contained in $L^1(\mu)$ and thus, it is easy to see the map $f \longrightarrow \int_X f\, d\mu$ defines a positive linear functional on $C_c(X)$. That is, we can naturally assign a positive linear functional on $C_c(X)$ to every locally finite Borel measure on X. The curious reader may ask if this assignment is unique and if the opposite scenario holds. That is, s/he may wonder if we may assign to each positive linear functional a unique locally finite Borel measure on X. This is indeed the case, and it is the principle result of the Riesz Representation Theorem for Radon measures.

Theorem 3.1 (Riesz Representation)**.** *If I is a positive linear functional on $C_c(X)$, then there exists a unique Radon measure μ on X such that*

$$I(f) = \int_X f \, d\mu \quad \textit{for each} \quad f \in C_c(X).$$

In addition, μ satisfies

$$\mu(U) = \sup\big\{I(f) \,\big|\, f \in C_c(X),\ f \prec U\big\} \quad \textit{for each open set} \quad U \subset X, \tag{3.1}$$

and

$$\mu(K) = \inf\big\{I(f) \,\big|\, f \in C_c(X),\ f \geq \chi_K\big\} \quad \textit{for each compact set} \quad K \subset X. \tag{3.2}$$

Proof. Suppose $I : C_c(X) \longrightarrow \mathbb{R}$ is a positive linear functional.

Step 1: We prove the uniqueness of the Radon measure μ.

If μ is a Radon measure such that $I(f) = \int_X f \, d\mu$ for all $f \in C_c(X)$, and $U \subset X$ is open, $f \prec U$ implies $I(f) \leq \int_X \chi_U \, d\mu = \mu(U)$. Thus,

$$\mu(U) \geq \sup\big\{I(f) \,|\, f \in C_c(X),\ f \prec U\big\}.$$

For a compact set $K \subset U$, Urysohn's lemma asserts there is a function $f \in C_c(X)$ such that $K \prec f \prec U$, and therefore $\mu(K) \leq \int_X f \, d\mu = I(f)$. By the inner regularity of Radon measures on open sets,

$$\mu(U) = \sup\big\{\mu(K) \,|\, K \subset U,\ K \text{ compact}\big\} \leq I(f).$$

This verifies μ satisfies (3.1). This also shows μ is determined uniquely by I on all open sets U in X. In fact, this holds on all Borel sets of X thanks to the outer regularity of μ on all Borel sets. This establishes the uniqueness of the Radon measure μ.

Step 2: We prove the existence of the Radon measure μ by constructing a suitable outer measure μ^* such that every open set in X is μ^*-measurable.

For open set $U \subset X$, we define

$$\mu(U) = \sup\big\{I(f) \,|\, f \in C_c(X),\ f \prec U\big\}. \tag{3.3}$$

and we define $\mu^* : 2^X \longrightarrow [0, +\infty]$ by

$$\mu^*(E) = \inf\big\{\mu(U) \,|\, U \supset E,\ U \text{ open}\big\} \tag{3.4}$$

and of course these coincide on all open sets, i.e. $\mu^*(U) = \mu(U)$ for each open set in X. This follows from the fact that $U \subset V$ implies $\mu(U) \leq \mu(V)$.

We now show μ^* is indeed an outer measure. Actually, we will show for any $E \subset X$,

$$\mu^*(E) = \inf\Big\{\sum_{i=1}^{\infty} \mu(U_j) \,\Big|\, U_j \text{ open},\ E \subset \bigcup_{i=1}^{\infty} U_j\Big\}, \tag{3.5}$$

which we know defines an outer measure by Proposition 1.1. To verify this claim, observe that for each open $U \supset E$,

$$\mu(U) \geq \inf\Big\{ \sum_{i=1}^{\infty} \mu(U_j) \,\Big|\, U_j \text{ open, } E \subset \bigcup_{i=1}^{\infty} U_j \Big\},$$

so taking the infimum of the left-hand side over all open $U \supset E$ shows

$$\mu^*(E) \geq \inf\Big\{ \sum_{i=1}^{\infty} \mu(U_j) \,\Big|\, U_j \text{ open, } E \subset \bigcup_{i=1}^{\infty} U_j \Big\}.$$

Conversely, to get the reverse inequality of the preceding estimate and thereby proving (3.5), we only need to show that if $\{U_j\}_{i=1}^{\infty}$ is a sequence of open sets and $U = \cup_{i=1}^{\infty} U_i$, then $\mu(U) \leq \sum_{i=1}^{\infty} \mu(U_i)$.

Indeed, if $U = \cup_{i=1}^{\infty} U_i$, $f \in C_c(X)$, and $f \prec U$, we set $K = supp(f) \subset U$. Since K is compact, $K \subset \cup_{i=1}^{n} U_i$ for some finite n. Thus, we have a partition of unity on K subordinate to $\{U_i\}_{i=1}^{\infty}$, i.e., there exist $g_1, g_2, \ldots, g_n$ of class $C_k(X)$ with $g_i \prec U_i$ and $\sum_{i=1}^{n} g_i = 1$ on K. Then

$$f = \sum_{i=1}^{n} f g_j \quad \text{and} \quad f g_i \prec U_i$$

and so

$$I(f) = \sum_{i=1}^{n} I(f g_i) \leq \sum_{i=1}^{n} \mu(U_i) \leq \sum_{i=1}^{\infty} \mu(U_i).$$

As f was chosen arbitrarily, the definition of (3.3) leads us to $\mu(U) \leq \sum_{i=1}^{\infty} \mu(U_i)$ as desired. Hence, μ^* defined in (3.5) is an outer measure.

Step 3: We claim every open set in X is μ^*-measurable.

More precisely, we need to show that, for any open set U and given any subset $E \subset X$ with $\mu^*(E) < +\infty$,

$$\mu^*(E) \geq \mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \cap U^c). \tag{3.6}$$

It suffices to prove (3.6) restricted to only open sets E. This is because for any subset $E \subset X$ with $\mu^*(E) < +\infty$ and for $\epsilon > 0$, we can find an open set $V \subset X$ such that $E \subset V$ and $\mu(V) < \mu^*(E) + \epsilon$. Then

$$\mu^*(E) + \epsilon > \mu(V) \geq \mu^*(V \cap U) + \mu^*(V \backslash U) \geq \mu^*(E \cap U) + \mu^*(E \backslash U)$$

and we recover (3.6) after sending $\epsilon \longrightarrow 0$. Thus, we will assume U and E are both open sets. Since their union $E \cup U$ is also open, given any $\epsilon > 0$, Urysohn's lemma ensures there exists $f \in C_c(X)$ such that $f \prec E \cup U$, and by definition of μ in (3.3), $I(f) > \mu(E \cup U) - \epsilon/2$.

Likewise, since $E\backslash supp(f)$ is also open, we can find $g \in C_c(X)$ such that $g \prec E\backslash supp(f)$ and $I(g) > \mu(E\backslash supp(f)) - \epsilon/2$. Observing that $f + g \prec E$, it follows that

$$\begin{aligned}\mu^*(E) = \mu(E) \geq I(f) + I(g) &> \mu(E \cap U) + \mu(E\backslash supp(f)) - \epsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E\backslash U) - \epsilon.\end{aligned}$$

Sending $\epsilon \longrightarrow 0$ in the last estimate leads to (3.6).

Step 4: Construct the Borel measure μ from the outer measure.

From the result of Step 3, Carathèdory's theorem (see Theorem 1.3) implies every Borel set is μ^*-measurable and that $\mu = \mu^*\big|_{\mathcal{B}_X}$ is a Borel measure. Moreover, our definition of (3.3) and (3.5) imply μ is outer regular on all Borel sets and satisfy (3.1).

Step 5: The Borel measure μ satisfies (3.2) and it is locally finite on compact sets.

Let $K \subset X$ be compact, $f \in C_c(X)$ with $f \geq \chi_K$, and let $U_\epsilon = \{x \in X \,|\, f(x) > 1 - \epsilon\}$. The set U_ϵ is open since f is continuous, $K \subset U_\epsilon$, and if $g \prec U_\epsilon$ we get $(1-\epsilon)^{-1} f - g \geq 0$ and thus $I(g) \leq (1-\epsilon)^{-1} I(f)$. This implies

$$\mu(K) \leq \mu(U_\epsilon) \leq (1-\epsilon)^{-1} I(f),$$

and we deduce $\mu(K) \leq I(f)$ after sending $\epsilon \longrightarrow 0$, i.e., we obtain

$$\mu(K) \leq \inf \big\{ I(f) \,|\, f \in C_c(X),\ f \geq \chi_K \big\} \ \text{ for each compact set } \ K \subset X. \tag{3.7}$$

It remains to obtain the reverse inequality of the last estimate. So, let $K \subset X$ be a compact set. Then, for any open set $U \supset K$, Urysohn's lemma ensures there exists $f \in C_c(X)$ such that $K \prec f \prec U$, and so $I(f) \leq \mu(U)$. Since μ is outer regular on K, we get

$$\begin{aligned}\mu(K) &= \inf\{\mu(U) \,|\, U \supset K,\ U \text{ open}\} \\ &\geq \inf \big\{ I(f) \,|\, f \in C_c(X),\ f \geq \chi_K \big\} \ \text{ for each compact set } \ K \subset X.\end{aligned}$$

This proves (3.2). In addition, (3.7) implies that $\mu(K) < +\infty$ for each compact subset $K \subset X$, i.e., μ is locally finite on compact sets.

Step 6: We prove $I(f) = \int_X f\, d\mu$ for each $f \in C_c(X)$.

Without loss of generality, we prove this for $f \in C_c(X)$ with $0 \leq f \leq 1$. For a positive integer N, define $K_j = \{x \in X \,|\, f(x) \geq jN^{-1}\}$ for $j = 1, 2, \ldots, N$, and set $K_0 = supp(f)$.

Define $f_1, f_2, \ldots, f_N \in C_c(X)$ by $f_j(x) = 0$ if $x \neq K_{j-1}$, $f_j(x) = f(x) - (j-1)N^{-1}$ if $x \in K_{j-1}\backslash K_j$, and $f_j(x) = N^{-1}$ if $x \in K_j$. That is,

$$f_j = \min\Big\{ \max\big\{ f - \frac{j-1}{N}, 0 \big\}, \frac{1}{N} \Big\}.$$

Then, each f_j is measurable and $N^{-1}\chi_{K_j} \leq f_j \leq N^{-1}\chi_{K_{j-1}}$ and so

$$\frac{1}{N}\mu(K_j) \leq \int_X f_j \, d\mu \leq \frac{1}{N}\mu(K_{j-1}). \tag{3.8}$$

Moreover, if U is an open set containing K_{j-1}, we have $Nf_j \prec U$. Thus, $I(f_j) \leq N^{-1}\mu(U)$. Then, by (3.2) and outer regularity, we deduce

$$\frac{1}{N}\mu(K_j) \leq I(f_j) \leq \frac{1}{N}\mu(K_{j-1}). \tag{3.9}$$

Indeed, we have $f = \sum_{j=1}^{N} f_j$, and therefore (3.8) and (3.9), respectively, imply

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq \int_X f\,d\mu \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j),$$

and

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leq I(f) \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j).$$

This further leads to

$$\left|I(f) - \int_X f\,d\mu\right| \leq \frac{1}{N}[\mu(K_0) - \mu(K_N)] \leq \frac{1}{N}\mu(supp(f)) \leq \frac{C}{N},$$

for some positive constant C independent of N. Sending $N \longrightarrow +\infty$, we arrive at

$$I(f) = \int_X f\,d\mu.$$

□

3.3 Further Properties of Radon Measures

This section reviews some additional properties of Radon measures. Recall that we define Radon measures to be inner regular on all open sets. It turns out inner regularity holds over all Borel sets provided that X is also σ-compact, i.e., there exists a collection $\{K_i\}_{i=1}^{\infty}$ of compact subsets of X such that $X = \cup_{i=1}^{n} K_i$. This will follow from the following result.

Theorem 3.2. *Every Radon measure is inner regular on all of its σ-finite sets.*

Proof. Let μ be a Radon measure on a measure space $(X, \mathcal{M}, \mu)$, and let E be any σ-finite set in X.

Case 1: If $\mu(E) < +\infty$, for arbitrary $\epsilon > 0$, we can pick an open set $U \supset E$ such that $\mu(U) < \mu(E) + \epsilon/2$ and a compact $F \subset U$ such that $\mu(F) > \mu(U) - \epsilon/2$.

Since $\mu(U \backslash E) < \epsilon$, we can choose an open set $V \supset U \backslash E$ such that $\mu(V) < \epsilon$. Set $K = F \backslash V$ so that K is compact, $K \subset E$, and

$$\mu(K) = \mu(F) - \mu(F \cap V) > \mu(E) - \epsilon/2 - \mu(V) > \mu(E) - \epsilon.$$

This shows μ is inner regular on E.

Case 2: If instead $\mu(E) = +\infty$, the we can find a sequence $\{E_i\}_{i=1}^{\infty}$ such that $E_1 \subset E_2 \subset E_3 \subset \ldots$, and $E = \cup_{i=1}^{\infty}$ such that $\mu(E_i) < +\infty$ for each i and $\mu(E_i) \longrightarrow +\infty$. Then, for each fixed natural number N, there exists i_0 such that $\mu(E_{i_0}) > N$ and hence, by applying the arguments in Case 1, there is a compact set $K \subset E_{i_0}$ with $\mu(K) > N$. Hence,

$$+\infty = \mu(E) = \sup\{\mu(K) \,\big|\, K \text{ is compact, } K \subset E\},$$

meaning μ is inner regular on E. □

Corollary 3.1. *Every σ-finite Radon measure is regular. Therefore, if X is σ-compact, then every Radon measure on X is regular.*

Remark. *We provide a measure space and a Radon measure where inner regularity does not hold on all Borel sets. Let $X = \mathbb{R} \times \mathbb{R}_d$, where $\mathbb{R}_d$ is the metric space $\mathbb{R}$ equipped with the discrete metric $d(x,y) = \chi_{\{x \neq y\}}$. If $f \in C(X)$, we define $f^y(x) = f(x,y)$; and if $E \subset X$, define $E^y = \{x \,|\, (x,y) \in E\}$.*

It follows that $f \in C_c(X)$ if and only if $f^y \in C_c(\mathbb{R})$ for all y and $f^y = 0$ for all but finitely-many y. Then $I(f) = \sum_{y \in \mathbb{R}} \int f(x,y)\,dx$ defines a positive linear functional on $C_c(X)$. Then, let μ be the unique Radon measure associated to I given by the Riesz Representation theorem of Theorem 3.1. Then, $\mu(E) = +\infty$ for any E such that $E^y \neq \emptyset$ for uncountably-many y.

Thus, if $E = \{0\} \times \mathbb{R}_d$, then $\mu(E) = +\infty$ but $\mu(K) = 0$ for all compact subsets $K \subset E$, i.e., μ is not inner regular on this set E.

Furthermore, the following reveals σ-compactness ensures locally finite Borel measures are indeed Radon.

Theorem 3.3. *Let X be a locally compact, Hausdorff space in which every open set is σ-compact (which is the case, e.g., if X is second countable). Then every Borel measure on X that is finite on compact sets is regular and hence Radon.*

We study some approximation theorems for Radon measures and characterize integration over Radon measures. The first is the density of test functions in the Lebesgue spaces with respect to Radon measures.

Theorem 3.4. *If μ is a Radon measure on X, then $C_c(X)$ is dense in $L^p(\mu)$, for $1 < p < \infty$.*

Proof. Recall that simple functions belonging to $L^p(\mu)$ are dense in $L^p(\mu)$. Thus, it suffices to show that for any Borel set E with $\mu(E) < +\infty$, χ_E can be approximated by functions in $C_c(X)$ with respect to the $L^p(\mu)$-norm.

Choose any $\epsilon > 0$. By Theorem 3.2, we can find a compact $K \subset E$ and an open set U such that $E \subset U$ and $\mu(U \backslash K) < \epsilon$. Further, by Urysohn's lemma, we can find $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_U$. This leads to

$$\|\chi_E - f\|_{L^p(\mu)}^p \leq \mu(U \backslash K) < \epsilon$$

and this completes the proof. □

For a certain class of non-negative functions, the so-called lower and upper semicontinuous functions, we have a nice description of their integrals against Radon measures. Let μ be a Radon measure on a measure space $(X, \mathcal{M}, \mu)$.

Definition. *For a topological space X, a function $f : X \longrightarrow (-\infty, +\infty]$ is said to be* **lower semicontinuous** *if the level set $\big\{x \in X \,:\, f(x) > t\big\}$ is open in X for each $t \in \mathbb{R}$. Likewise, a function $f : X \longrightarrow [-\infty, +\infty)$ is said to be* **upper semicontinuous** *if the level set $\big\{x \in X \,:\, f(x) < t\big\}$ is an open set in X for each $t \in \mathbb{R}$.*

Proposition 3.2. *Let (X, τ) be a topological space.*

(a) If U is open in X, then χ_U is lower semicontinuous.

(b) If f and g are lower semicontinuous and $c \geq 0$, then cf and $f + g$ are lower semicontinuous.

(c) If $\mathcal{G}$ is a family of lower semicontinuous functions and

$$f(x) = \sup_{g \in \mathcal{G}} g(x),$$

then f is lower semicontinuous.

(d) If, in addition, X is locally compact and Hausdorff and f is a non-negative, lower semicontinuous function, then

$$f(x) = \sup\big\{g(x) \,\big|\, g \in C_c(X),\, 0 \leq g \leq f\big\}.$$

Theorem 3.5. *Let $\mathcal{G}$ be a family of non-negative, lower semicontinuous functions on a locally compact Hausdorff space X that is directed by $\leq$, i.e., for every $g_1, g_2 \in \mathcal{G}$, there exists $g \in mathcalG$ such thtat $g_1 \leq g$ and $g_2 \leq g$. Let $f = \sup\{g \,|\, g \in \mathcal{G}\}$. If μ is any Radon measure on X, then*

$$\int_X f\, d\mu = \sup_{g \in \mathcal{G}} \int_X g\, d\mu.$$

Part (d) of Proposition 3.2 and Theorem 3.5 imply the following.

Corollary 3.2. *If μ is a Radon measure and f is a non-negative, lower semicontinuous function, then*

$$\int_X f\, d\mu = \sup\Big\{ \int_X g\, d\mu \,\Big|\, g \in C_c(X),\, 0 \leq g \leq f\Big\}.$$

We have the following characterization for integration of non-negative Borel measurable functions.

Theorem 3.6. *If μ is a Radon measure and f is a non-negative, Borel measurable function, then*

$$\int_X f\,d\mu = \inf\Big\{ \int_X g\,d\mu \,\Big|\, f \leq g,\ g \text{ is lower semicontinuous}\Big\}.$$

If $\{x \in X \mid f(x) > 0\}$ is σ-finite, then

$$\int_X f\,d\mu = \sup\Big\{ \int_X g\,d\mu \,\Big|\, 0 \leq g \leq f,\ \text{and } g \text{ is upper semicontinuous}\Big\}.$$

3.4 Signed and complex Radon measures

This section introduces signed Radon measures and the complex Radon measures, which forms a proper normed linear space. This space actually gives an equivalent characterization of the dual space of $C_0(X)$, the linear space of continuous linear functions on $C_0(X)$, thereby extending the Riesz Representation theorem.

3.4.1 The dual space of $C_0(X)$.

Let X be a locally compact Hausdorff space. Then recall $C_0(X)$ denotes the uniform closure of $C_c(X)$. Hence, if μ is a Radon measure on X, then the linear functional $I(f) = \int_X f\,d\mu$ can be extended continuously to $C_0(X)$ if and only if it bounded, i.e., continuous, with respect to the uniform norm.

In view of this, the basic inequality of Lemma 2.2, and because

$$\mu(X) = \sup\Big\{ \int_X f\,d\mu \,\Big|\, f \in C_c(X),\ 0 \leq f \leq 1\Big\}$$

due to (3.1) in Theorem 3.1, then I is bounded so long as $\mu(X) < +\infty$, in which case $\mu(X)$ is equal to the operator norm of I. In other words, the positive and bounded linear functionals on $C_0(X)$ are characterized by integration against finite Radon measures on X. One main objective of this section is to further refine this result and give a complete description of the dual space $C_0(X)^*$.

We shall require the following Jordan-type decomposition for real-valued linear functions on $C_0(X, \mathbb{R})$.

Lemma 3.1. *If $I \in C_0(X, \mathbb{R})^*$, then there exist positive linear functionals $I^\pm \in C_0(X, \mathbb{R})^*$ such that $I = I^+ - I^-$.*

Definition. *A* **signed Radon measure** *is a signed Borel measure whose positive and negative variations are Radon, and a* **complex Radon measure** *is a complex Borel measure whose real and imaginary parts are signed Radon measures.*

Remark. *In view of Theorem 3.3 and since complex measures are bounded, every complex Borel measure is Radon on locally compact Hausdorff spaces that are second countable.*

Definition. *The denote the collection of all complex Radon measures on X by $M(X)$, and for $\mu \in M(X)$, we define*

$$\|\mu\| = |\mu|(X), \tag{3.10}$$

where $|\mu|$ is the total variation of μ.

Indeed, $M(X)$ defines a normed linear space.

Theorem 3.7. *If μ is a complex Borel measure on X, then μ is Radon if and only if $|\mu|$ is Radon. Moreover, $M(X)$ is a linear space and $\mu \mapsto \|\mu\|$ defines a norm on $M(X)$.*

We have the following refinement of the Riesz Representation theorem.

Theorem 3.8. *Let X be a locally compact Hausdorff space, and for $\mu \in M(X)$ and $f \in C_0(X)$, let $I_\mu(f) = \int_X f\,d\mu$. Then the mapping $\mu \mapsto I_\mu$ is an isometric isomorphism from $M(X)$ onto $C_0(X)^*$.*

Bibliography

[1] S. Axler. *Measure, Integration & Real Analysis.* Graduate Texts in Mathematics. Springer, 2020.

[2] R. G. Bartle. *The Elements of Integration and Lebesgue Measure*. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999.

[3] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications. Second Edition.* Pure and Applied Mathematics (New York). A Wiley-Interscience Publication, 1999.

[4] W. Rudin. *Real & Complex Analysis. Third Edition.* McGraw-Hill Book Co., New York, 1987.