Abstract Measure And Integration

1.1 $\sigma$ Algebras

This chapter is on the basics of measure theory and integration. A measure is a real valued mapping from some subset of the power set of a given set which has values in $[0, \infty]$. We will see that many apparently different things can be considered as measures and also that whenever we are in such a measure space defined below, there is an integral defined. By discussing this in terms of axioms and in a very abstract setting, we unify many topics into one general theory. For example, it will turn out that sums are included as an integral of this sort. So is the usual integral as well as things which are often thought of as being in between sums and integrals.

Let $\Omega$ be a set and let $\mathcal{F}$ be a collection of subsets of $\Omega$ satisfying

$$\emptyset \in \mathcal{F}, \; \Omega \in \mathcal{F}, \tag{1.1}$$

$$E \in \mathcal{F} \text{ implies } E^C \equiv \Omega \setminus E \in \mathcal{F},$$

If $\{E_n\}_{n=1}^\infty \subseteq \mathcal{F}$, then $\cup_{n=1}^\infty E_n \in \mathcal{F}. \tag{1.2}$

**Definition 1.1.1** A collection of subsets of a set, $\Omega$, satisfying Formulas 1.1-1.2 is called a $\sigma$ algebra.

As an example, let $\Omega$ be any set and let $\mathcal{F} = \mathcal{P}(\Omega)$, the set of all subsets of $\Omega$ (power set). This obviously satisfies Formulas 1.1-1.2.

**Lemma 1.1.2** Let $\mathcal{C}$ be a set whose elements are $\sigma$ algebras of subsets of $\Omega$. Then $\cap \mathcal{C}$ is a $\sigma$ algebra also.

Be sure to verify this lemma. It follows immediately from the above definitions but it is important for you to check the details.

**Example 1.1.3** Let $\tau$ denote the collection of all open sets in $\mathbb{R}^n$ and let $\sigma(\tau) \equiv \text{intersection of all } \sigma \text{ algebras that contain } \tau$. $\sigma(\tau)$ is called the $\sigma$ algebra of Borel sets. In general, if we have a collection of sets, $\Sigma$, we denote by $\sigma(\Sigma)$ the smallest $\sigma$ algebra which contains $\Sigma$.

This is a very important $\sigma$ algebra and it will be referred to frequently as the Borel sets. Attempts to describe a typical Borel set are more trouble than they are worth and it is not easy to do so. Rather, one uses the definition just given in the example. Note, however, that all countable intersections of open sets and countable unions of closed sets are Borel sets. Such sets are called $G_\delta$ and $F_\sigma$ respectively.
**Definition 1.1.4** Let $\mathcal{F}$ be a $\sigma$ algebra of sets of $\Omega$ and let $\mu : \mathcal{F} \to [0, \infty]$. We call $\mu$ a measure if

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i) \quad (1.3)$$

whenever the $E_i$ are disjoint sets of $\mathcal{F}$. The triple, $(\Omega, \mathcal{F}, \mu)$ is called a measure space and the elements of $\mathcal{F}$ are called the measurable sets. We say $(\Omega, \mathcal{F}, \mu)$ is a finite measure space when $\mu(\Omega) < \infty$.

The following theorem is the basis for most of what is done in the theory of measure and integration. It is a very simple result which follows directly from the above definition.

**Theorem 1.1.5** Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of measurable sets in a measure space $(\Omega, \mathcal{F}, \mu)$. Then if \(\cdots \subseteq E_n \subseteq E_{n+1} \subseteq \cdots\),

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n) \quad (1.4)$$

and if \(\cdots \supseteq E_n \supseteq E_{n+1} \supseteq \cdots\) and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n). \quad (1.5)$$

Stated more succinctly, $E_k \uparrow E$ implies $\mu(E_k) \uparrow \mu(E)$ and $E_k \downarrow E$ with $\mu(E_1) < \infty$ implies $\mu(E_k) \downarrow \mu(E)$.

**Proof:** First note that $\bigcap_{k=1}^{\infty} E_k = (\bigcup_{k=1}^{\infty} E_k^c)^c \in \mathcal{F}$ so $\bigcap_{k=1}^{\infty} E_k$ is measurable. Also note that for $A$ and $B$ sets of $\mathcal{F}$, $A \setminus B \equiv (A^c \cup B)^c \in \mathcal{F}$. To show 1.4, note that 1.4 is obviously true if $\mu(E_k) = \infty$ for any $k$. Therefore, assume $\mu(E_k) < \infty$ for all $k$. Thus

$$\mu(E_{k+1} \setminus E_k) + \mu(E_k) = \mu(E_{k+1})$$

and so

$$\mu(E_{k+1} \setminus E_k) = \mu(E_{k+1}) - \mu(E_k).$$

Also,

$$\bigcup_{k=1}^{\infty} E_k = E_1 \cup \bigcup_{k=1}^{\infty} (E_{k+1} \setminus E_k)$$

and the sets in the above union are disjoint. Hence by 1.3,

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \mu(E_1) + \sum_{k=1}^{\infty} \mu(E_{k+1} \setminus E_k) = \mu(E_1)$$

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

$$= \mu(E_1) + \lim_{n \to \infty} \sum_{k=1}^{n} \mu(E_{k+1}) - \mu(E_k) = \lim_{n \to \infty} \mu(E_{n+1}).$$

This shows part 1.4.
1.1. $\sigma$ ALGEBRAS

To verify 1.5,

$$\mu(E_1) = \mu(\cap_{i=1}^{\infty} E_i) + \mu(E_1 \setminus \cap_{i=1}^{\infty} E_i)$$

since $\mu(E_1) < \infty$, it follows $\mu(\cap_{i=1}^{\infty} E_i) < \infty$. Also, $E_1 \setminus \cap_{i=1}^{n} E_i \uparrow E_1 \setminus \cap_{i=1}^{\infty} E_i$ and so by 1.4,

$$\mu(E_1) - \mu(\cap_{i=1}^{n} E_i) = \mu(E_1 \setminus \cap_{i=1}^{n} E_i) = \lim_{n \to \infty} \mu(E_1 \setminus \cap_{i=1}^{n} E_i)$$

$$= \mu(E_1) - \lim_{n \to \infty} \mu(\cap_{i=1}^{n} E_i) = \mu(E_1) - \lim_{n \to \infty} \mu(E_n),$$

Hence, subtracting $\mu(E_1)$ from both sides,

$$\lim_{n \to \infty} \mu(E_n) = \mu(\cap_{i=1}^{\infty} E_i).$$

This proves the theorem.

We will sometimes need to consider functions which can take the value $+\infty$. You should think of $+\infty$, usually refered to as $\infty$ as something out at the right end of the real line and its only importance is the notion of sequences converging to it. We say $x_n \to \infty$ exactly when for all $l \in \mathbb{R}$, there exists $N$ such that if $n \geq N$, then

$$x_n > l.$$  

This is what it means for a sequence to converge to $\infty$. Don’t think of $\infty$ as a number. It is just a convenient symbol which allows the consideration of some limit operations more simply. Similar considerations apply to $-\infty$ but we really don’t care about letting functions assume this value so it will be left off.

**Lemma 1.1.6** Let $f : \Omega \to (-\infty, \infty]$ where $\mathcal{F}$ is a $\sigma$ algebra of subsets of $\Omega$. Then the following are equivalent.

$$f^{-1}((d, \infty]) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}((-\infty, d)) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}([d, \infty]) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}((-\infty, d]) \in \mathcal{F} \text{ for all finite } d,$$

$$f^{-1}((a, b)) \in \mathcal{F} \text{ for all } a < b, -\infty < a < b < \infty.$$

**Proof:** First note that the first and the third are equivalent. To see this, observe

$$f^{-1}([d, \infty]) = \cap_{n=1}^{\infty} f^{-1}((d - 1/n, \infty]),$$

and so if the first condition holds, then so does the third.

$$f^{-1}((a, b)) = \cup_{n=1}^{\infty} f^{-1}([d + 1/n, \infty]),$$

and so if the third condition holds, so does the first.
Similarly, the second and fourth conditions are equivalent. Now
\[ f^{-1}((-\infty, d]) = (f^{-1}((d, \infty]))^c \]
so the first and fourth conditions are equivalent. Thus the first four conditions are equivalent and if any of them hold, then for \(-\infty < a < b < \infty\),
\[ f^{-1}((a, b)) = f^{-1}((-\infty, b)) \cap f^{-1}((a, \infty)) \in \mathcal{F}. \]
Finally, if the last condition holds,
\[ f^{-1}([d, \infty)) = (\bigcup_{k=1}^\infty f^{-1}((-k + d, d)))^c \in \mathcal{F} \]
and so the third condition holds. Therefore, all five conditions are equivalent. This proves the lemma.

This lemma allows for the following definition of a measurable function having values in \((-\infty, \infty]\).

**Definition 1.1.7** Let \((\Omega, \mathcal{F}, \mu)\) be a measure space and let \(f : \Omega \to (-\infty, \infty]\). Then \(f\) is said to be measurable if any of the equivalent conditions of Lemma 1.1.6 hold. When the \(\sigma\) algebra, \(\mathcal{F}\) equals the Borel \(\sigma\) algebra, \(\mathcal{B}\), the function is called Borel measurable.

**Theorem 1.1.8** Let \(f_n\) and \(f\) be functions mapping \(\Omega\) to \((-\infty, \infty]\) where \(\mathcal{F}\) is a \(\sigma\) algebra of measurable sets of \(\Omega\). Then if \(f_n\) is measurable, and \(f(\omega) = \lim_{n \to \infty} f_n(\omega)\), it follows that \(f\) is also measurable. (Pointwise limits of measurable functions are measurable.)

**Proof:** We show \(f^{-1}((a, b)) \in \mathcal{F}\). Let \(V_m \equiv (a + \frac{1}{m}, b - \frac{1}{m})\) and \(\overline{V}_m = [a + \frac{1}{m}, b - \frac{1}{m}]\). Then for all \(m\), \(V_m \subseteq (a, b)\) and
\[(a, b) = \bigcup_{m=1}^\infty V_m = \bigcup_{m=1}^\infty \overline{V}_m.\]
Note that \(V_m \neq \emptyset\) for all \(m\) large enough. Since \(f\) is the pointwise limit of \(f_n\),
\[ f^{-1}(V_m) \subseteq \{\omega : f_k(\omega) \in V_m \text{ for all } k \text{ large enough} \} \subseteq f^{-1}(\overline{V}_m). \]
You should note that the expression in the middle is of the form
\[ \bigcup_{m=1}^\infty \cap_{k=n}^\infty f_k^{-1}(V_m). \]
Therefore,
\[ f^{-1}((a, b)) = \bigcup_{m=1}^\infty f^{-1}(V_m) \subseteq \bigcup_{m=1}^\infty \bigcap_{k=m}^\infty f_k^{-1}(V_m) \]
\[ \subseteq \bigcup_{m=1}^\infty f^{-1}(\overline{V}_m) = f^{-1}((a, b)). \]
It follows \(f^{-1}((a, b)) \in \mathcal{F}\) because it equals the expression in the middle which is measurable. This shows \(f\) is measurable.

**Theorem 1.1.9** Let \(B\) consist of open cubes of the form
\[ Q_x \equiv \prod_{i=1}^n (x_i - \delta, x_i + \delta) \]
where \(\delta\) is a positive rational number and \(x \in \mathbb{Q}^n\). Then every open set in \(\mathbb{R}^n\) can be written as a countable union of open cubes from \(B\). Furthermore, \(B\) is a countable set.
1.1. \( \sigma \) ALGEBRAS

**Proof:** Let \( U \) be an open set and let \( y \in U \). Since \( U \) is open, \( B(y, r) \subseteq U \) for some \( r > 0 \) and we can assume \( r/\sqrt{n} \in \mathbb{Q} \). Let
\[
x \in B\left(y, \frac{r}{10\sqrt{n}}\right) \cap \mathbb{Q}^n
\]
and consider the cube, \( Q_x \in \mathcal{B} \) defined by
\[
Q_x = \prod_{i=1}^{n} (x_i - \delta, x_i + \delta)
\]
where \( \delta = r/4\sqrt{n} \). The following picture is roughly illustrative of what is taking place.

![Diagram showing the relationship between \( B(y, r) \), \( Q_x \), and \( y \).]

Then the diameter of \( Q_x \) equals
\[
\left( \frac{n}{2\sqrt{n}} \right)^{1/2} = \frac{r}{2}
\]
and so, if \( z \in Q_x \), then
\[
|z - y| \leq |z - x| + |x - y| < \frac{r}{2} + \frac{r}{2} = r.
\]
Consequently, \( Q_x \subseteq U \). Now also,
\[
\left( \sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2} < \frac{r}{10\sqrt{n}}
\]
and so it follows that for each \( i \),
\[
|x_i - y_i| < \frac{r}{4\sqrt{n}}
\]
since otherwise the above inequality would not hold. Therefore, \( y \in Q_x \subseteq U \). Now let \( \mathcal{B}_U \) denote those sets of \( \mathcal{B} \) which are contained in \( U \). Then \( \cup \mathcal{B}_U = U \).

To see \( \mathcal{B} \) is countable, note there are countably many choices for \( x \) and countably many choices for \( \delta \). This proves the theorem.

Recall that \( g : \mathbb{R}^n \to \mathbb{R} \) is continuous means \( g^{-1} \) (open set) = an open set. In particular \( g^{-1} ((a, b)) \) must be an open set.

**Theorem 1.1.10** Let \( f_i : \Omega \to \mathbb{R} \) for \( i = 1, \cdots, n \) be measurable functions and let \( g : \mathbb{R}^n \to \mathbb{R} \) be continuous where \( f \equiv (f_1 \cdots f_n)^T \). Then \( g \circ f \) is a measurable function from \( \Omega \) to \( \mathbb{R} \).
Proof: We need to show
\[(g \circ f)^{-1}((a, b)) \in \mathcal{F}.
\]
Now \((g \circ f)^{-1}((a, b)) = f^{-1}(g^{-1}((a, b)))\) and since \(g\) is continuous, it follows that \(g^{-1}((a, b))\) is an open set which we denote as \(U\) for convenience. Now by Theorem 1.1.9 above, it follows there are countably many open cubes, \(\{Q_k\}\) such that
\[U = \bigcup_{k=1}^{\infty} Q_k\]
where each \(Q_k\) is a cube of the form
\[Q_k = \prod_{i=1}^{n} (x_i - \delta, x_i + \delta).
\]
Now
\[f^{-1} \left( \prod_{i=1}^{n} (x_i - \delta, x_i + \delta) \right) = \bigcap_{i=1}^{n} f_i^{-1}((x_i - \delta, x_i + \delta)) \in \mathcal{F}
\]
and so
\[(g \circ f)^{-1}((a, b)) = f^{-1}(g^{-1}((a, b))) = f^{-1}(U) = f^{-1}(\bigcup_{k=1}^{\infty} Q_k) = \bigcup_{k=1}^{\infty} f^{-1}(Q_k) \in \mathcal{F}.
\]
This proves the theorem.

Corollary 1.1.11 Sums, products, and linear combinations of measurable functions are measurable.

Proof: To see the product of two measurable functions is measurable, let \(g(x, y) = xy\), a continuous function defined on \(\mathbb{R}^2\). Thus if you have two measurable functions, \(f_1\) and \(f_2\) defined on \(\Omega\),
\[g \circ (f_1, f_2)(\omega) = f_1(\omega) f_2(\omega)
\]
and so \(\omega \rightarrow f_1(\omega) f_2(\omega)\) is measurable. Similarly you can show the sum of two measurable functions is measurable by considering \(g(x, y) = x + y\) and you can show a linear combination of two measurable functions is measurable by considering \(g(x, y) = ax + by\). More than two functions can also be considered as well.

The message of this corollary is that starting with measurable real valued functions you can combine them in pretty much any way you want and you end up with a measurable function.

Here is some notation which will be used whenever convenient.

Definition 1.1.12 Let \(f : \Omega \rightarrow [-\infty, \infty]\). We define
\[\preceq f \equiv \{ \omega \in \Omega : f(\omega) > \alpha \} \equiv f^{-1}((\alpha, \infty])\]
with obvious modifications for the symbols \(\preceq f, [\alpha \leq f, [\alpha \geq f, [\alpha \geq f \geq \beta], etc.

Theorem 1.1.13 (Egoroff) Let \((\Omega, \mathcal{F}, \mu)\) be a finite measure space
\[(\mu(\Omega) < \infty)\]
and let $f_n, f$ be complex valued measurable functions such that Re $f_n$, Im $f_n$, Re $f$, and Im $f$ are all measurable. Then

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

for all $\omega \notin E$ where $\mu(E) = 0$. Then for every $\varepsilon > 0$, there exists a set,

$$F \supseteq E, \mu(F) < \varepsilon,$$

such that $f_n$ converges uniformly to $f$ on $F^c$.

**Proof:** Let $E_{km} = \{\omega \in E^c : |f_n(\omega) - f(\omega)| \geq 1/m \text{ for some } n > k\}$. Note that

$$|f_n(\omega) - f(\omega)| = \sqrt{(\text{Re } f_n(\omega) - \text{Re } f(\omega))^2 + (\text{Im } f_n(\omega) - \text{Im } f(\omega))^2}$$

and so, By Theorem 1.1.10,

$$\left| |f_n - f| \geq \frac{1}{m} \right|$$

is measurable. Hence $E_{km}$ is measurable because

$$E_{km} = \bigcup_{n=k+1}^{\infty} \left| |f_n - f| \geq \frac{1}{m} \right|.$$

For fixed $m$, $\cap_{k=1}^{\infty} E_{km} = \emptyset$ because we are given that $f_n$ converges to $f$ off $E$. Therefore, if $\omega \notin E$ there exists $k$ such that if $n > k$, $|f_n(\omega) - f(\omega)| < \frac{1}{m}$ which means $\omega \notin E_{km}$. Therefore, this set has measure zero. Note also that

$$E_{km} \supseteq E_{(k+1)m}.$$

Since $\mu(E_{1m}) < \infty$, we can apply Theorem 1.1.5 on Page 2 to conclude

$$0 = \mu(\cap_{k=1}^{\infty} E_{km}) = \lim_{k \to \infty} \mu(E_{km}).$$

Let $k(m)$ be chosen such that $\mu(E_{k(m)m}) < \varepsilon 2^{-m}$ and let

$$F = E \cup \bigcup_{m=1}^{\infty} E_{k(m)m}.$$

Then $\mu(F) < \varepsilon$ because

$$\mu(F) \leq \mu(E) + \sum_{m=1}^{\infty} \mu(E_{k(m)m})$$

$$< 0 + \sum_{m=1}^{\infty} \varepsilon 2^{-m} = \varepsilon$$

Now let $\eta > 0$ be given and pick $m_0$ such that $m_0^{-1} < \eta$. If $\omega \in F^c$, then

$$\omega \in \bigcap_{m=1}^{\infty} E_{k(m)m}^c.$$
Hence $\omega \in F_{k(m_{0})m_{0}}^{c}$, so

$$|f_{n}(\omega) - f(\omega)| < 1/m_{0} < \eta$$

for all $n > k(m_{0})$. This holds for all $\omega \in F^{c}$ and so $f_{n}$ converges uniformly to $f$ on $F^{c}$. This proves the theorem.

We conclude this section with a comment about notation.

**Definition 1.1.14** We say that something happens for $\mu$ a.e. $\omega$ and say $\mu$ almost everywhere if there exists a set $E$ with $\mu(E) = 0$ and the thing takes place for all $\omega \notin E$. Thus $f(\omega) = g(\omega)$ a.e. if $f(\omega) = g(\omega)$ for all $\omega \notin E$ where $\mu(E) = 0$. We also say a measure space, $(\Omega, \mathcal{F}, \mu)$, is $\sigma$ finite if there exist measurable sets, $\Omega_{n}$ such that $\mu(\Omega_{n}) < \infty$ and $\Omega = \bigcup_{n=1}^{\infty} \Omega_{n}$.

### 1.2 Exercises

1. Let $\Omega = \mathbb{N} = \{1, 2, \cdots \}$. Let $\mathcal{F} = \mathcal{P}(\mathbb{N})$ and let $\mu(S)$ = number of elements in $S$. Thus $\mu(\{1\}) = 1 = \mu(\{2\})$, $\mu(\{1, 2\}) = 2$, etc. Show $(\Omega, \mathcal{F}, \mu)$ is a measure space. It is called counting measure. What functions are measurable in this case?

2. Let $\Omega$ be any uncountable set and let $\mathcal{F} = \{A \subseteq \Omega : \text{either } A \text{ or } A^{c} \text{ is countable}\}$. Let $\mu(A) = 1$ if $A$ is uncountable and $\mu(A) = 0$ if $A$ is countable. Show $(\Omega, \mathcal{F}, \mu)$ is a measure space. This is a well known bad example.

3. Let $\mathcal{F}$ be a $\sigma$ algebra of subsets of $\Omega$ and suppose $\mathcal{F}$ has infinitely many elements. Show that $\mathcal{F}$ is uncountable. **Hint:** You might try to show there exists a countable sequence of disjoint sets of $\mathcal{F}$, $\{A_{i}\}$. It might be easiest to verify this by contradiction if it doesn’t exist rather than a direct construction. Once this has been done, you can define a map, $\theta$, from $\mathcal{P}(\mathbb{N})$ into $\mathcal{F}$ which is one to one by $\theta(S) = \bigcup_{i \in S} A_{i}$. Then argue $\mathcal{P}(\mathbb{N})$ is uncountable and so $\mathcal{F}$ is also uncountable.

4. Prove Lemma 1.1.2.

5. We say $g$ is Borel measurable if whenever $U$ is open, $g^{-1}(U)$ is Borel. Let $f : \Omega \to \mathbb{R}^{n}$ and let $g : \mathbb{R}^{n} \to \mathbb{R}$ and $\mathcal{F}$ is a $\sigma$ algebra of sets of $\Omega$. Suppose $f$ is measurable and $g$ is Borel measurable. Show $g \circ f$ is measurable. To say $g$ is Borel measurable means $g^{-1}(\text{open set}) = (\text{Borel set})$ where a Borel set is one of those sets in the smallest $\sigma$ algebra containing the open sets of $\mathbb{R}^{n}$. See Lemma 1.1.2. **Hint:** You should show, using Theorem 1.1.9 that $f^{-1}(\text{open set}) \in \mathcal{F}$. Now let

$$S \equiv \{E \subseteq \mathbb{R}^{n} : f^{-1}(E) \in \mathcal{F}\}$$

By what you just showed, $S$ contains the open sets. Now verify $S$ is a $\sigma$ algebra. Argue that from the definition of the Borel sets, it follows $S$ contains the Borel sets.

6. Let $(\Omega, \mathcal{F})$ be a measure space and suppose $f : \Omega \to \mathbb{C}$. Then $f$ is said to be measurable if

$$f^{-1}(\text{open set}) \in \mathcal{F}.$$ 

Show $f$ is measurable if and only if $\text{Re } f$ and $\text{Im } f$ are measurable real-valued functions. Thus we may as well define a complex valued
function to be measurable if the real and imaginary parts are measurable. **Hint:** Argue that 
\( f^{-1}(((a, b) + i(c, d))) = (\text{Re } f)^{-1}((a, b)) \cap (\text{Im } f)^{-1}((c, d)). \) Then use Theorem 1.1.9 to verify that if \( \text{Re } f \) and \( \text{Im } f \) are measurable, it follows \( f \) is. Conversely, argue that \( (\text{Re } f)^{-1}((a, b)) = f^{-1}((a, b) + i\mathbb{R}) \) with a similar formula holding for \( \text{Im } f \).

7. Let \((\Omega, \mathcal{F}, \mu)\) be a measure space. Define \( \mathcal{P}: \mathcal{P}(\Omega) \to [0, \infty] \) by

\[
\mathcal{P}(A) = \inf \{ \mu(B) : B \supseteq A, B \in \mathcal{F} \}.
\]

Show \( \mathcal{P} \) satisfies

\[
\mathcal{P}(\emptyset) = 0, \text{ if } A \subseteq B, \mathcal{P}(A) \leq \mathcal{P}(B), \\
\mathcal{P}(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mathcal{P}(A_i), \mu(A) = \mathcal{P}(A) \text{ if } A \in \mathcal{F}.
\]

If \( \mathcal{P} \) satisfies these conditions, it is called an outer measure. This shows every measure determines an outer measure on the power set. We return to outer measures later.

8. Let \( \{E_i\} \) be a sequence of measurable sets with the property that

\[
\sum_{i=1}^{\infty} \mu(E_i) < \infty.
\]

Let \( S = \{ \omega \in \Omega \text{ such that } \omega \in E_i \text{ for infinitely many values of } i \} \). Show \( \mu(S) = 0 \) and \( S \) is measurable. This is part of the Borel Cantelli lemma. **Hint:** Write \( S \) in terms of intersections and unions. Something is in \( S \) means that for every \( n \) there exists \( k > n \) such that it is in \( E_k \). Remember the tail of a convergent series is small.

9. \( \uparrow \) Let \( f_n, f \) be measurable functions. We say that \( f_n \) converges in measure if

\[
\lim_{n \to \infty} \mu(\Omega : |f(x) - f_n(x)| \geq \epsilon) = 0
\]

for each fixed \( \epsilon > 0 \). Prove the theorem of F. Riesz. If \( f_n \) converges to \( f \) in measure, then there exists a subsequence \( \{f_{n_k}\} \) which converges to \( f \) a.e. **Hint:** Choose \( n_1 \) such that

\[
\mu(x : |f(x) - f_{n_1}(x)| \geq 1) < 1/2.
\]

Choose \( n_2 > n_1 \) such that

\[
\mu(x : |f(x) - f_{n_2}(x)| \geq 1/2) < 1/2^2,
\]

\( n_3 > n_2 \) such that

\[
\mu(x : |f(x) - f_{n_3}(x)| \geq 1/3) < 1/2^3,
\]

etc. Now consider what it means for \( f_{n_k}(x) \) to fail to converge to \( f(x) \). Then use Problem 8.
1.3 The Abstract Lebesgue Integral

In this section we develop the Lebesgue integral along with the major convergence theorems which are the reason for studying the Lebesgue integral. In all that follows $\mu$ will be a measure defined on a $\sigma$ algebra $\mathcal{F}$ of subsets of $\Omega$. We always define $0 \cdot \infty = 0$. This is a meaningless expression and so we are free to define it as we please and a little thought will soon demonstrate that this is the right definition in the context of measure theory. To see this, consider the zero function defined on $\mathbb{R}$. What do we want the integral of this function to be? Obviously, by an analogy with the Riemann integral, we would want this to equal zero. Formally, it is zero times the length of the set or infinity. Therefore, we adopt this convention in what follows. The following notation will be used.

For a set $E$,

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

This is called the characteristic function of $E$. Sometimes this is called the indicator function which I think is better terminology since the term characteristic function has another meaning. Note that this “indicates” whether a point, $\omega$ is contained in $E$. It is exactly when the function has the value 1.

**Lemma 1.3.1** Let $f(a, b) \in [-\infty, \infty]$ for $a \in A$ and $b \in B$ where $A, B$ are sets. Then

$$
\sup_{a \in A} \sup_{b \in B} f(a, b) = \sup_{b \in B} \sup_{a \in A} f(a, b).
$$

**Proof:** Note that for all $a, b$, $f(a, b) \leq \sup_{a \in A} f(a, b)$ and therefore, for all $a$, $\sup_{b \in B} f(a, b) \leq \sup_{b \in B} \sup_{a \in A} f(a, b)$. Therefore,

$$
\sup_{a \in A} \sup_{b \in B} f(a, b) \leq \sup_{b \in B} \sup_{a \in A} f(a, b).
$$

Repeating the same argument interchanging $a$ and $b$, gives the conclusion of the lemma.

**Lemma 1.3.2** If $\{A_n\}$ is an increasing sequence in $[-\infty, \infty]$, then $\sup A_n = \lim_{n \to \infty} A_n$.

The following lemma is useful also and this is a good place to put it. First we say $\{b_j\}_{j=1}^{\infty}$ is an enumeration of the $a_{ij}$ is

$$
\bigcup_{j=1}^{\infty} \{b_j\} = \bigcup_{i,j} \{a_{ij}\}.
$$

In other words, we list the countable set, $\{a_{ij}\}_{i,j=1}^{\infty}$ as $b_1, b_2, \ldots$

**Lemma 1.3.3** Let $a_{ij} \geq 0$. Then $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$. Also if $\{b_j\}_{j=1}^{\infty}$ is any enumeration of the $a_{ij}$, then $\sum_{j=1}^{\infty} b_j = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}$.

**Proof:** First we note there is no trouble in defining these sums because the $a_{ij}$ are all nonnegative. If a sum diverges, it only diverges to $\infty$ and we write $\infty$ for the answer.

$$
\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \geq \sup_{n} \sum_{j=1}^{\infty} \sum_{i=1}^{n} a_{ij} = \sup_{m} \lim_{n \to \infty} \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}
$$
\[ = \sup_n \lim_{m \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} = \sup_n \sum_{i=1}^{n} \sum_{j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}. \]  

Interchanging the \( i \) and \( j \) in the above argument the first part of the lemma is proved.

Finally, note that for all \( p \),

\[ \sum_{j=1}^{p} b_j \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \]

and so \( \sum_{j=1}^{\infty} b_j \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \). Now let \( m, n > 1 \) be given. Then

\[ \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \leq \sum_{i=1}^{\infty} \sum_{j=1}^{p} b_j \]

where \( p \) is chosen large enough that \( \{b_1, \cdots, b_p\} \supseteq \{a_{ij} : i \leq m \text{ and } j \leq n\} \).

Therefore, since such a \( p \) exists for any choice of \( m, n \), it follows that for any \( m, n \),

\[ \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \leq \sum_{j=1}^{\infty} b_j. \]

Therefore, taking the limit as \( n \to \infty \),

\[ \sum_{i=1}^{m} \sum_{j=1}^{\infty} a_{ij} \leq \sum_{j=1}^{\infty} b_j \]

and finally, taking the limit as \( m \to \infty \),

\[ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \leq \sum_{j=1}^{\infty} b_j \]

proving the lemma.

The following picture illustrates the idea we wish to use in defining the Lebesgue integral to be like the area under the curve.

![Diagram of the Lebesgue integral](image)

You can see that by following the procedure illustrated in the picture and letting \( h \) get smaller, you would expect to obtain better approximations to the area under the curve\(^1\) although all these approximations would likely be too small. Therefore, define

\(^1\)Note the difference between this picture and the one usually drawn in calculus courses where the little rectangles are upright rather than on their sides. This illustrates a fundamental philosophical difference between the Riemann and the Lebesgue integrals. With the Riemann integral we measure intervals. With the Lebesgue integral, we measure inverse images of intervals.
\[
\int f \, d\mu \equiv \sup_{h > 0} \sum_{i=1}^{\infty} h \mu ([ih < f])
\]

**Lemma 1.3.4** The following inequality holds,
\[
\sum_{i=1}^{\infty} h \mu ([ih < f]) \leq \sum_{i=1}^{\infty} \frac{h}{2^i} \mu \left( \left\lfloor \frac{h}{2} < f \right\rfloor \right).
\]
Also, it suffices to consider only \( h \) smaller than a given positive number in the above definition of the integral.

**Proof:**
Let \( N \in \mathbb{N} \).
\[
\sum_{i=1}^{2N} \frac{h}{2^i} \mu \left( \left\lfloor \frac{h}{2} < f \right\rfloor \right) = \sum_{i=1}^{2N} \frac{h}{2^i} \mu ([ih < 2f])
\]
\[
= \sum_{i=1}^{N} \frac{h}{2^i} \mu ([2i-1] h < 2f]) + \sum_{i=1}^{N} \frac{h}{2^i} \mu ([2i] h < 2f])
\]
\[
= \sum_{i=1}^{N} \frac{h}{2^i} \mu \left( \left\lfloor \frac{2i-1}{2} h < f \right\rfloor \right) + \sum_{i=1}^{N} \frac{h}{2^i} \mu ([ih < f])
\]
\[
\geq \sum_{i=1}^{N} \frac{h}{2^i} \mu ([ih < f]) + \sum_{i=1}^{N} \frac{h}{2^i} \mu ([ih < f]) = \sum_{i=1}^{N} h \mu ([ih < f])
\]

Now letting \( N \to \infty \) yields the claim of the lemma.

**Definition 1.3.5** A function, \( s \), is called simple if it is a measurable real valued function and has only finitely many values. These values will never be \( \pm \infty \). Thus a simple function is one which may be written in the form
\[
s(\omega) = \sum_{i=1}^{n} c_i \chi_{E_i}(\omega)
\]
where the sets, \( E_i \) are disjoint and measurable. The set, \( E_i \) is where \( s \) takes the value \( c_i \).

Note that by taking the union of some of the \( E_i \) in the above definition, we can assume that the numbers, \( c_i \) are the distinct values of \( s \). Simple functions are important because it will turn out to be very easy to take their integrals as shown in the following lemma.

**Lemma 1.3.6** Let \( s(\omega) = \sum_{i=1}^{p} a_i \chi_{E_i}(\omega) \) be a nonnegative simple function with the \( a_i \) the distinct values of \( s \). Then
\[
\int s \, d\mu = \sum_{i=1}^{p} a_i \mu(E_i).
\]
Also, for any nonnegative measurable function, \( f \), if \( \lambda \geq 0 \), then
\[
\int \lambda f \, d\mu = \lambda \int f \, d\mu.
\]
1.3. THE ABSTRACT LEBESGUE INTEGRAL

**Proof:** Without loss of generality, we can assume \( 0 < a_1 < a_2 < \cdots < a_p \) and that \( \mu (E_i) < \infty \). Let \( a_0 \equiv 0 \). By Lemma 1.3.4 we can assume \( 0 < 2h < \min \{ a_i - a_{i-1}\}^p_{i=1} \). Let \( k_i (h) \) be a positive integer defined for each \( i = 1, 2, \ldots, p \) by

\[
k_i (h) r_i = a_i
\]

where \( 0 \leq r_i < h \). Then

\[
\sum_{i=1}^{\infty} h \mu ([ih < s]) = \sum_{i=1}^{k_1 (h)} h \mu ([ih < s]) + \sum_{i=k_1 (h)+1}^{k_2 (h)} h \mu ([ih < s]) + \\
\cdots + \sum_{i=k_{p-1} (h)+1}^{k_p (h)} h \mu ([ih < s])
\]

\[
h k_1 (h) \sum_{i=1}^{p} \mu (E_i) - f_1 (h) + h (k_2 (h) - k_1 (h)) \sum_{i=2}^{p} \mu (E_i) - f_2 (h) + \\
\cdots + h (k_p (h) - k_{p-1} (h)) \mu (E_p) - f_p (h)
\]

where \( f_j (h) \) either equals \( 0 \) or \( h \mu (E_j) \), depending on whether \( r_j = 0 \). Therefore, this equals

\[
\sum_{j=1}^{p} (a_j - r_j - (a_{j-1} - r_{j-1})) \sum_{i=j}^{p} \mu (E_i) - g (h)
\]

where \( \lim_{h \to 0} g (h) = 0 \) and \( g (h) \geq 0 \). Interchanging the order of the summation,

\[
\sum_{i=1}^{p} \mu (E_i) \sum_{j=1}^{i} (a_j - r_j - (a_{j-1} - r_{j-1})) - g (h)
\]

\[
= \sum_{i=1}^{p} a_i \mu (E_i) - \sum_{i=1}^{p} r_i \mu (E_i) - g (h).
\]

Taking the sup yields 1.7 because \( r_i < h \).

To verify 1.8 we note the formula is obvious if \( \lambda = 0 \) because then \([ih < \lambda f] = \emptyset \) for all \( i > 0 \). Assume \( \lambda > 0 \). Then

\[
\int \lambda f d\mu \equiv \sup_{h > 0} \sum_{i=1}^{\infty} h \mu ([ih < \lambda f])
\]

\[
= \sup_{h > 0} \sum_{i=1}^{\infty} h \mu ([ih / \lambda < f])
\]

\[
= \sup_{h > 0} \lambda \sum_{i=1}^{\infty} (h / \lambda) \mu ([ih / \lambda < f])
\]

\[
= \lambda \int f d\mu.
\]

This proves the lemma.
Lemma 1.3.7  If we write the nonnegative simple function,

\[ s(\omega) = \sum_{i=1}^{n} c_i X_{E_i}(\omega) \]

where the \( c_i \) are not necessarily distinct but the \( E_i \) are disjoint, it follows that

\[ \int s = \sum_{i=1}^{n} c_i \mu(E_i). \]

Proof: Let the values of \( s \) be \( \{a_1, \ldots, a_m\} \). Therefore, since the \( E_i \) are disjoint, we must have each \( a_i \) equal to one of the \( c_j \). Let \( A_i \equiv \cup\{E_j : c_j = a_i\} \). Then from Lemma 1.3.6 it follows that

\[ \int s = \sum_{i=1}^{m} a_i \mu(A_i) = \sum_{i=1}^{m} a_i \sum_{\{j : c_j = a_i\}} \mu(E_j) \]

\[ = \sum_{i=1}^{m} \sum_{\{j : c_j = a_i\}} c_j \mu(E_j) = \sum_{i=1}^{n} c_i \mu(E_i). \]

This proves the lemma.

Note that \( \int s \) could equal \( +\infty \) if \( \mu(A_k) = \infty \) and \( a_k > 0 \) for some \( k \), but \( \int s \) is well defined because \( s \geq 0 \) and we use the convention that \( 0 \cdot \infty = 0 \).

Lemma 1.3.8  If \( a, b \geq 0 \) and if \( s \) and \( t \) are nonnegative simple functions, then

\[ \int as + bt = a \int s + b \int t. \]

Proof: Let

\[ s(\omega) = \sum_{i=1}^{n} \alpha_i X_{A_i}(\omega), \quad t(\omega) = \sum_{i=1}^{n} \beta_j X_{B_j}(\omega) \]

where the \( \alpha_i \) are the distinct values of \( s \) and the \( \beta_j \) are the distinct values of \( t \). Clearly \( as + bt \) is a nonnegative simple function because it is measurable and has finitely many values. Also,

\[ (as + bt)(\omega) = \sum_{j=1}^{m} \sum_{i=1}^{n} (a\alpha_i + b\beta_j) X_{A_i \cap B_j}(\omega) \]

where the sets \( A_i \cap B_j \) are disjoint. By Lemma 1.3.7,

\[ \int as + bt = \sum_{j=1}^{m} \sum_{i=1}^{n} (a\alpha_i + b\beta_j) \mu(A_i \cap B_j) \]

\[ = a \sum_{i=1}^{n} \alpha_i \mu(A_i) + b \sum_{j=1}^{m} \beta_j \mu(B_j) \]

\[ = a \int s + b \int t. \]

This proves the lemma.

There is a fundamental theorem about the relationship of simple functions to measurable functions given in the next theorem.
1.3. THE ABSTRACT LABESQUE INTEGRAL

**Theorem 1.3.9** Let \( f \geq 0 \) be measurable. Then there exists a sequence of simple functions \( \{s_n\} \) satisfying

\[
0 \leq s_n(\omega)
\]

\[
\cdots s_n(\omega) \leq s_{n+1}(\omega) \cdots
\]

\[
f(\omega) = \lim_{n \to \infty} s_n(\omega) \text{ for all } \omega \in \Omega.
\]

If \( f \) is bounded the convergence is actually uniform.

**Proof:** Letting \( I \equiv \{ \omega : f(\omega) = \infty \} \), define

\[
t_n(\omega) = \sum_{k=0}^{\infty} \frac{1}{n} \chi_{k/n \leq f < (k+1)/n}(\omega) + n \chi_f(\omega).
\]

Then \( t_n(\omega) \leq f(\omega) \) for all \( \omega \) and \( \lim_{n \to \infty} t_n(\omega) = f(\omega) \) for all \( \omega \). This is because \( t_n(\omega) = n \) for \( \omega \in I \) and if \( f(\omega) \in [0, 2^{n+1}) \), then

\[
0 \leq f(\omega) - t_n(\omega) \leq \frac{1}{n}
\]

Thus whenever \( \omega \notin I \), the above inequality will hold for all \( n \) large enough. Let

\[
s_1 = t_1, \ s_2 = \max(t_1, t_2), \ s_3 = \max(t_1, t_2, t_3), \ldots
\]

Then the sequence \( \{s_n\} \) satisfies Formulas 1.9-1.10.

To verify the last claim, note that in this case the term \( n \chi_f(\omega) \) is not present. Therefore, for all \( n \) large enough we have 1.11 holding for all \( \omega \). Thus the convergence is uniform. This proves the theorem.

Now we give a short proof of the monotone convergence theorem based on Lemma 1.3.1 and this definition of the integral. It is nothing but a computation.

**Theorem 1.3.10** (Monotone Convergence Theorem) Let \( f \) have values in \([0, \infty]\) and suppose \( \{f_n\} \) is a sequence of nonnegative measurable functions having values in \([0, \infty]\) and satisfying

\[
\lim_{n \to \infty} f_n(\omega) = f(\omega) \text{ for each } \omega.
\]

\[
\cdots f_n(\omega) \leq f_{n+1}(\omega) \cdots
\]

Then \( f \) is measurable and

\[
\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu.
\]
Proof: From Lemmas 1.3.1 and 1.3.2,

\[ \int f \, d\mu = \sup_{k > 0} \sum_{i=1}^{\infty} h_{\mu}([i \in f]) \]

\[ = \sup_{k > 0} \sum_{i=1}^{k} h_{\mu}([i \in f]) \]

\[ = \sup_{k > 0} \sup_{m} \sum_{i=1}^{k} h_{\mu}([i \in f_m]) \]

\[ = \sup_{m} \sum_{i=1}^{\infty} h_{\mu}([i \in f_m]) \]

\[ = \sup_{m} \int f_m \, d\mu \]

\[ = \lim_{m \to \infty} \int f_m \, d\mu. \]

The third equality follows from the observation that

\[ \lim_{m \to \infty} \mu([i \in f_m]) = \mu([i \in f]) \]

which follows from Theorem 1.1.5 since the sets, \([i \in f_m]\) are increasing in \(m\) and their union equals \([i \in f]\). This proves the theorem.

Next we show that the integral wants to be linear.

Theorem 1.3.11 Let \(f, g\) be nonnegative measurable functions and let \(a, b\) be nonnegative numbers. Then

\[ \int (af + bg) \, d\mu = a \int f \, d\mu + b \int g \, d\mu. \tag{1.12} \]

Proof: We use Theorem 1.3.9 on Page 15 to obtain increasing sequences of nonnegative simple functions, \(s_n \to f\) and \(t_n \to g\). Then by the monotone convergence theorem and Lemma 1.3.8,

\[ \int (af + bg) \, d\mu = \lim_{n \to \infty} \int a s_n + b t_n \, d\mu \]

\[ = \lim_{n \to \infty} \left(a \int s_n \, d\mu + b \int t_n \, d\mu\right) \]

\[ = a \int f \, d\mu + b \int g \, d\mu. \]

1.4 The Space \(L^1\)

Definition 1.4.1 Suppose \(f\) has complex values. Then we say that \(f\) is measurable if both \(\text{Re} \, f\) and \(\text{Im} \, f\) are measurable real valued functions.

Definition 1.4.2 \(L^1(\Omega)\) is the space of complex valued measurable functions, \(f\), satisfying

\[ \int |f(\omega)| \, d\mu < \infty. \]

We also write the symbol, \(\|f\|_{L^1}\), to denote \(\int |f(\omega)| \, d\mu\).
1.4. THE SPACE $L^1$

Note that if $f : \Omega \to \mathbb{C}$ is measurable, then by Theorem 1.1.10, $|f| : \Omega \to \mathbb{R}$ is also measurable because

$$|f(\omega)| = \sqrt{(\text{Re } f(\omega))^2 + (\text{Im } f(\omega))^2}$$

and both $\text{Re } f$ and $\text{Im } f$ are measurable.

**Definition 1.4.3** If $u$ is real-valued,

$$u^+ \equiv \max(u, 0), \quad u^- \equiv -\min(u, 0).$$

The following lemma follows immediately from the above definition.

**Lemma 1.4.4** $u^+$ and $u^-$ are both nonnegative and

$$u = u^+ - u^-, \quad |u| = u^+ + u^-.$$

**Definition 1.4.5** Let $f = u + iv$ where $u, v$ are real-valued. Suppose $f \in L^1(\Omega)$. Then we define

$$\int f \, d\mu \equiv \int u^+ \, d\mu - \int u^- \, d\mu + i \int v^+ \, d\mu - i \int v^- \, d\mu.$$ 

Note that all this is well defined because $\int |f| \, d\mu < \infty$ and so

$$\int u^+ \, d\mu, \int u^- \, d\mu, \int v^+ \, d\mu, \int v^- \, d\mu$$

are all finite. The next theorem shows the integral is linear on $L^1(\Omega)$.

**Theorem 1.4.6** $L^1(\Omega)$ is a complex vector space and if $a, b \in \mathbb{C}$ and $f, g \in L^1(\Omega)$,

then

$$\int a f + b g \, d\mu = a \int f \, d\mu + b \int g \, d\mu. \quad (1.13)$$

**Proof:** First suppose $f, g$ are real-valued and in $L^1(\Omega)$. (Note we are not allowing these functions to take the value $+\infty$.) The definition implies that for $h$ a real valued function,

$$h^+ = 2^{-1}(h + |h|), \quad h^- = 2^{-1}(|h| - h)$$

Consequently, the following formula follows immediately.

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

Since none of the functions have values equal to $\pm \infty$, this implies

$$f^+ + g^+ + (f + g)^- = (f + g)^+ + f^- + g^- \quad (1.14)$$

From Lemma 1.3.11,

$$\int f^+ \, d\mu + \int g^+ \, d\mu + \int (f + g)^- \, d\mu = \int f^- \, d\mu + \int g^- \, d\mu + \int (f + g)^+ \, d\mu \quad (1.15)$$
Now \( f(g) \leq |f| + |g| \) and similarly, \( f(g)^+ \leq |f| + |g| \) as are all the other integrals. Since all integrals are finite,

\[
\int (f + g)^+ \, d\mu - \int (f + g)^- \, d\mu = \int f^+ \, d\mu + \int g^+ \, d\mu - \left( \int f^- \, d\mu + \int g^- \, d\mu \right)
\]

The definition and the above implies

\[
\int (f + g) \, d\mu \quad \equiv \quad \int (f + g)^+ \, d\mu - \int (f + g)^- \, d\mu \quad (1.16)
\]

\[
= \int f^+ \, d\mu + \int g^+ \, d\mu - \left( \int f^- \, d\mu + \int g^- \, d\mu \right)
\]

\[
= \int f \, d\mu + \int g \, d\mu.
\]

We need to be able to factor out constants also. Suppose that \( c \) is a real constant and \( f \) is real-valued. Note

\[
(cf)^- = -cf^+ \text{ if } c < 0, \quad (cf)^- = cf^- \text{ if } c \geq 0.
\]

\[
(cf)^+ = -cf^- \text{ if } c < 0, \quad (cf)^+ = cf^+ \text{ if } c \geq 0.
\]

If \( c < 0 \), we use the above and Lemma 1.3.11 to write

\[
\int cf \, d\mu = \int (cf)^+ \, d\mu - \int (cf)^- \, d\mu = \int f^- \, d\mu + c \int f^+ \, d\mu = c \int f \, d\mu.
\]

Similarly, if \( c \geq 0 \),

\[
\int cf \, d\mu \quad \equiv \quad \int (cf)^+ \, d\mu - \int (cf)^- \, d\mu = c \int f^+ \, d\mu - c \int f^- \, d\mu = c \int f \, d\mu.
\]

This shows 1.13 holds if \( f, g, a, \) and \( b \) are all real-valued. To conclude, let \( a = \alpha + i\beta \), \( f = u + iv \) and use the preceding,

\[
\int af \, d\mu = \int (\alpha + i\beta)(u + iv) \, d\mu = \int (\alpha u - \beta v) + i(\beta u + \alpha v) \, d\mu = \alpha \int ud\mu - \beta \int vd\mu + i\beta \int ud\mu + i\alpha \int vd\mu = (\alpha + i\beta)(\int ud\mu + i \int vd\mu) = a \int f \, d\mu.
\]

Thus 1.13 holds whenever \( f, g, a, \) and \( b \) are complex valued.

If \( f, g \in L^1(\Omega) \) and \( a, b \) are constants, \(|af + bg|\) is measurable by Theorem 1.1.10 applied to \( f(Re) \) and \( f(Im) \). Also,

\[
\int |af + bg| \, d\mu \leq \int |a| |f| + |b| |g| \, d\mu < \infty
\]

thanks to the assumption that \( f, g \in L^1(\Omega) \). Therefore, \( L^1(\Omega) \) is a vector space.

Sometimes the limit of a sequence does not exist. There are two more general notions known as lim sup and lim inf which do always exist in some sense. These notions are dependent on the following lemma.
1.4. THE SPACE $L^1$

**Lemma 1.4.7** Let $\{a_n\}$ be an increasing (decreasing) sequence in $[-\infty, \infty]$. Then $\lim_{n \to \infty} a_n$ exists.

**Proof:** Suppose first $\{a_n\}$ is increasing. Recall this means $a_n \leq a_{n+1}$ for all $n$. If the sequence is bounded above, then it has a least upper bound and so $a_n \to a$ where $a$ is its least upper bound. If the sequence is not bounded above, then for every $l \in \mathbb{R}$, it follows $l$ is not an upper bound and so eventually, $a_n > l$. But this is what we mean when we say $a_n \to \infty$. The situation for decreasing sequences is completely similar.

Now take any sequence, $\{a_n\} \subseteq [-\infty, \infty]$ and consider the sequence $\{A_n\}$ where $A_n \equiv \inf \{a_k : k \geq n\}$. Then as $n$ increases, the set of numbers whose inf is being taken is getting smaller. Therefore, $A_n$ is an increasing sequence and so it must converge. Similarly, if $B_n \equiv \sup \{a_k : k \geq n\}$, it follows $B_n$ is decreasing and so $\{B_n\}$ also must converge. With this preparation, the following definition can be given.

**Definition 1.4.8** Let $\{a_n\}$ be a sequence of points in $[-\infty, \infty]$. Then we define

$$\liminf_{n \to \infty} a_n \equiv \liminf_{n \to \infty} \{a_k : k \geq n\}$$

and

$$\limsup_{n \to \infty} a_n \equiv \limsup_{n \to \infty} \{a_k : k \geq n\}$$

In the case of functions having values in $[-\infty, \infty]$, $(\liminf_{n \to \infty} f_n)(\omega) \equiv \liminf_{n \to \infty} (f_n(\omega))$. A similar definition applies to $\limsup_{n \to \infty} f_n$.

**Lemma 1.4.9** Let $\{a_n\}$ be a sequence in $[-\infty, \infty]$. Then $\lim_{n \to \infty} a_n$ exists if and only if $\liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n$ and in this case, the limit equals the common value of these two numbers.

**Proof:** Suppose first $\lim_{n \to \infty} a_n = a \in \mathbb{R}$. Then, letting $\varepsilon > 0$ be given, $a_n \in (a-\varepsilon, a+\varepsilon)$ for all $n$ large enough, say $n \geq N$. Therefore, both $\inf \{a_k : k \geq n\}$ and $\sup \{a_k : k \geq n\}$ are contained in $[a-\varepsilon, a+\varepsilon]$ whenever $n \geq N$. It follows $\limsup_{n \to \infty} a_n$ and $\liminf_{n \to \infty} a_n$ are both in $[a-\varepsilon, a+\varepsilon]$, showing

$$\left| \liminf_{n \to \infty} a_n - \limsup_{n \to \infty} a_n \right| < 2\varepsilon.$$  

Since $\varepsilon$ is arbitrary, the two must be equal and they both must equal $a$. Next suppose $\lim_{n \to \infty} a_n = \infty$. Then if $l \in \mathbb{R}$, there exists $N$ such that for $n \geq N$,

$$l \leq a_n$$

and therefore, for such $n$,

$$l \leq \inf \{a_k : k \geq n\} \leq \sup \{a_k : k \geq n\}$$

and this shows, since $l$ is arbitrary that

$$\liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n = \infty.$$  

The case for $-\infty$ is similar.
Conversely, suppose \( \liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n = a \). Suppose first that \( a \in \mathbb{R} \). Then, letting \( \varepsilon > 0 \) be given, there exists \( N \) such that if \( n \geq N \),
\[
\sup\{a_k : k \geq n\} - \inf\{a_k : k \geq n\} < \varepsilon
\]
therefore, if \( k, m > N \), and \( a_k > a_m \),
\[
|a_k - a_m| = a_k - a_m \leq \sup\{a_k : k \geq n\} - \inf\{a_k : k \geq n\} < \varepsilon
\]
showing that \( \{a_n\} \) is a Cauchy sequence. Therefore, it converges to \( a \in \mathbb{R} \) and as in the first part, the \( \liminf \) and \( \limsup \) both equal \( a \). If \( \liminf_{n \to \infty} a_n = \limsup_{n \to \infty} a_n = \infty \), then given \( l \in \mathbb{R} \), there exists \( N \) such that for \( n \geq N \),
\[
\inf_{n \geq N} a_n > l.
\]
Therefore, \( \lim_{n \to \infty} a_n = \infty \). The case for \(-\infty\) is similar. This proves the lemma.

The next theorem, known as Fatou’s lemma is another important theorem which justifies the use of the Lebesgue integral.

**Theorem 1.4.10 (Fatou’s lemma)** Let \( f_n \) be a nonnegative measurable function with values in \([0, \infty]\). Let \( g(\omega) = \liminf_{n \to \infty} f_n(\omega) \). Then \( g \) is measurable and
\[
\int gd\mu \leq \liminf_{n \to \infty} \int f_n d\mu.
\]

**Proof:** Let \( g_n(\omega) = \inf\{f_k(\omega) : k \geq n\} \). Then
\[
g_n^{-1}([a, \infty]) = \bigcap_{k=n}^{\infty} f_k^{-1}([a, \infty]) \in \mathcal{F}.
\]
Thus \( g_n \) is measurable by Lemma 1.1.6 on Page 3. Also \( g(\omega) = \lim_{n \to \infty} g_n(\omega) \) so \( g \) is measurable because it is the pointwise limit of measurable functions. Now the functions \( g_n \) form an increasing sequence of nonnegative measurable functions so the monotone convergence theorem applies. This yields
\[
\int gd\mu = \lim_{n \to \infty} \int g_n d\mu \leq \liminf_{n \to \infty} \int f_n d\mu.
\]
The last inequality holding because
\[
\int g_n d\mu \leq \int f_n d\mu.
\]
(We don’t know \( \lim_{n \to \infty} \int f_n d\mu \) exists.) This proves the Theorem.

The following is a fundamental property of the complex numbers which we will have occasion to use repeatedly.

**Lemma 1.4.11** If \( z \in \mathbb{C} \) there exists \( \alpha \in \mathbb{C} \) such that \( \alpha z = |z| \) and \( |\alpha| = 1 \).

**Proof:** Suppose \( z = x + iy \). If \( z = 0 \), then any complex number having modulus 1 will work. Therefore, assume \( z \neq 0 \). Let \( \alpha = \frac{z-iy}{\sqrt{x^2+y^2}} \) in this case.

**Theorem 1.4.12 (Triangle inequality)** Let \( f \in L^1(\Omega) \). Then
\[
|\int f d\mu| \leq \int |f| d\mu.
\]
1.4. THE SPACE $L^1$

**Proof:** $\int f d\mu \in \mathbb{C}$ so by Lemma 1.4.11 there exists $\alpha \in \mathbb{C}$, $|\alpha| = 1$ such that

$$|\int f d\mu| = \alpha \int f d\mu = \alpha \int f d\mu.$$

Hence

$$|\int f d\mu| = \int \alpha f d\mu = \int \text{Re}(\alpha f) d\mu + i \int \text{Im}(\alpha f) d\mu$$

$$= \int \text{Re}(\alpha f) d\mu = \int (\text{Re}(\alpha f))^+ d\mu - \int (\text{Re}(\alpha f))^- d\mu$$

$$\leq \int (\text{Re}(\alpha f))^+ + (\text{Re}(\alpha f))^- d\mu \leq \int |\text{Re}(\alpha f)| d\mu$$

$$\leq \int |\alpha f| d\mu = \int |f| d\mu$$

which proves the theorem.

**Theorem 1.4.13 (Dominated Convergence theorem)** Let $f_n \in L^1(\Omega)$ and suppose

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

and there exists a measurable function $g$, with values in $[0, \infty]$, such that

$$|f_n(\omega)| \leq g(\omega)$$

and $\int g(\omega) d\mu < \infty$.

Then $f \in L^1(\Omega)$ and

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

**Proof:** $f$ is measurable by Theorem 1.1.8. Since $|f| \leq g$, it follows that

$$f \in L^1(\Omega)$$

and $|f - f_n| \leq 2g$.

By Fatou’s lemma (Theorem 1.4.10),

$$\int 2g d\mu \leq \liminf_{n \to \infty} \int 2g - |f - f_n| d\mu$$

$$= \int 2g d\mu - \limsup_{n \to \infty} \int |f - f_n| d\mu.$$

Subtracting $\int 2g d\mu,$

$$0 \leq -\limsup_{n \to \infty} \int |f - f_n| d\mu.$$ 

Hence

$$0 \geq \limsup_{n \to \infty} \left( \int |f - f_n| d\mu \right) \geq \limsup_{n \to \infty} |\int f d\mu - \int f_n d\mu|$$

$$\geq \liminf_{n \to \infty} |\int f d\mu - \int f_n d\mu| \geq 0.$$

This proves the theorem by Lemma 1.4.9 on Page 19 because the $\limsup$ and $\liminf$ are equal.

---

Note that, since $g$ is allowed to have the value $\infty$, it is not known that $g \in L^1(\Omega)$. 
Definition 1.4.14 Let $E$ be a measurable subset of $\Omega$.

$$\int_E f \, d\mu \equiv \int f \chi_E \, d\mu.$$  

Also we may refer to $L^1(E)$. The $\sigma$ algebra in this case is just

$$\{ E \cap A : A \in \mathcal{F} \}$$

and the measure is $\mu$ restricted to this smaller $\sigma$ algebra. Clearly, if $f \in L^1(\Omega)$, then

$$f \chi_E \in L^1(E)$$

and if $f \in L^1(E)$, then letting $\tilde{f}$ be the 0 extension of $f$ off of $E$, we see that $\tilde{f} \in L^1(\Omega)$.

### 1.5 Vitali Convergence Theorem

In this section we consider a remarkable convergence theorem which, in the case of finite measure spaces turns out to be better than the dominated convergence theorem.

Definition 1.5.1 Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $\mathcal{S} \subseteq L^1(\Omega)$. We say that $\mathcal{S}$ is uniformly integrable if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $f \in \mathcal{S}$

$$|\int_E f \, d\mu| < \varepsilon \text{ whenever } \mu(E) < \delta.$$  

Lemma 1.5.2 If $\mathcal{S}$ is uniformly integrable, then $|\mathcal{S}| \equiv \{|f| : f \in \mathcal{S}\}$ is uniformly integrable. Also $\mathcal{S}$ is uniformly integrable if $\mathcal{S}$ is finite.

**Proof:** Let $\varepsilon > 0$ be given and suppose $\mathcal{S}$ is uniformly integrable. First suppose the functions are real valued. Let $\delta$ be such that if $\mu(E) < \delta$, then

$$\left| \int_E f \, d\mu \right| < \frac{\varepsilon}{2}$$

for all $f \in \mathcal{S}$. Let $\mu(E) < \delta$. Then if $f \in \mathcal{S}$,

$$\int_E |f| \, d\mu \leq \int_{E \cap \{|f| \leq 0\}} |f| \, d\mu + \int_{E \cap \{|f| > 0\}} f \, d\mu$$

$$= \left| \int_{E \cap \{|f| \leq 0\}} f \, d\mu \right| + \left| \int_{E \cap \{|f| > 0\}} f \, d\mu \right|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$  

In general, if $\mathcal{S}$ is a uniformly integrable set of complex valued functions, the inequalities,

$$\left| \int_E \text{Re} \ f \, d\mu \right| \leq \left| \int_E f \, d\mu \right|, \left| \int_E \text{Im} \ f \, d\mu \right| \leq \left| \int_E f \, d\mu \right|,$$

imply $\text{Re} \ \mathcal{S} \equiv \{ \text{Re} \ f : f \in \mathcal{S} \}$ and $\text{Im} \ \mathcal{S} \equiv \{ \text{Im} \ f : f \in \mathcal{S} \}$ are also uniformly integrable. Therefore, applying the above result for real valued
functions to these sets of functions, we verify that $|\mathcal{S}|$ is uniformly integrable also.

For the last part, is suffices to verify a single function in $L^1(\Omega)$ is uniformly integrable. To do so, note that from the dominated convergence theorem,

$$
\lim_{R \to \infty} \int_{|f| > R} |f| \, d\mu = 0.
$$

Let $\varepsilon > 0$ be given and choose $R$ large enough that $\int_{|f| > R} |f| \, d\mu < \frac{\varepsilon}{2}$. Now let $\mu(E) < \frac{\varepsilon}{2R}$. Then

$$
\int_E |f| \, d\mu = \int_{E \cap \{|f| \leq R\}} |f| \, d\mu + \int_{E \cap \{|f| > R\}} |f| \, d\mu
\quad < \quad R \mu(E) + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
$$

This proves the lemma.

The following theorem is Vitali’s convergence theorem.

**Theorem 1.5.3** Let $\{f_n\}$ be a uniformly integrable set of complex valued functions, $\mu(\Omega) < \infty$, and $f_n(x) \to f(x)$ a.e. where $f$ is a measurable complex valued function. Then $f \in L^1(\Omega)$ and

$$
\lim_{n \to \infty} \int_{\Omega} |f_n - f| \, d\mu = 0. 
\tag{1.17}
$$

**Proof:** First we show that $f \in L^1(\Omega)$. By uniform integrability, there exists $\delta > 0$ such that if $\mu(E) < \delta$, then

$$
\int_E |f_n| \, d\mu < 1
$$

for all $n$. By Egoroff’s theorem, there exists a set, $E$ of measure less than $\delta$ such that on $E^C$, $\{f_n\}$ converges uniformly. Therefore, if we pick $p$ large enough, and let $n > p$,

$$
\int_{E^c} |f_p - f_n| \, d\mu < 1
$$

which implies

$$
\int_{E^c} |f_n| \, d\mu < 1 + \int_{\Omega} |f_p| \, d\mu.
$$

Then since there are only finitely many functions, $f_n$ with $n \leq p$, we have the existence of a constant, $M_1$ such that for all $n$,

$$
\int_{E^c} |f_n| \, d\mu < M_1.
$$

But also, we have

$$
\int_{\Omega} |f| \, d\mu = \int_{E^c} |f| \, d\mu + \int_{E} |f| \, d\mu
\quad \leq \quad M_1 + 1 = M.
$$

Therefore, by Fatou’s lemma,

$$
\int_{\Omega} |f| \, d\mu \leq \lim_{n \to \infty} \inf \int_{\Omega} |f_n| \, d\mu \leq M,
$$
showing that $f \in L^1$ as hoped.

Now $\mathcal{G} \cup \{f\}$ is uniformly integrable so there exists $\delta_1 > 0$ such that if $\mu(E) < \delta_1$, then $\int_E |g| \, d\mu < \varepsilon / 3$ for all $g \in \mathcal{G} \cup \{f\}$. By Egoroff’s theorem, there exists a set, $F$ with $\mu(F) < \delta_1$ such that $f_n$ converges uniformly to $f$ on $F^c$. Therefore, there exists $N$ such that if $n > N$, then

$$\int_{F^c} |f - f_n| \, d\mu < \frac{\varepsilon}{3}.$$  

It follows that for $n > N$,

$$\int_{\Omega} |f - f_n| \, d\mu \leq \int_{F^c} |f - f_n| \, d\mu + \int_F |f| \, d\mu + \int_F |f_n| \, d\mu < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$  

which verifies 1.17.

### 1.6 Exercises

1. Let $\Omega = \mathbb{N} = \{1, 2, \ldots\}$ and $\mu(S) =$ number of elements in $S$. If

$$f : \Omega \to \mathbb{C}$$

what do we mean by $\int f \, d\mu$? Which functions are in $L^1(\Omega)$? Which functions are measurable?

2. Show that for $f \geq 0$ and measurable, $\int f \, d\mu \equiv \lim_{h \to 0^+} \sum_{i=1}^{\infty} h \mu(\{i h < f\})$.

3. For the measure space of Problem 1, give an example of a sequence of nonnegative measurable functions $\{f_n\}$ converging pointwise to a function $f$, such that inequality is obtained in Fatou’s lemma.

4. Fill in all the details of the proof of Lemma 1.5.2.

5. Let $\sum_{i=1}^{n} c_i \chi_{E_i}(\omega) = s(\omega)$ be a nonnegative simple function for which the $c_i$ are the distinct nonzero values. Show with the aid of the monotone convergence theorem that the two definitions of the Lebesgue integral given in the chapter are equivalent.

6. Suppose $(\Omega, \mu)$ is a finite measure space and $\mathcal{G} \subseteq L^1(\Omega)$. Show $\mathcal{G}$ is uniformly integrable and bounded in $L^1(\Omega)$ if there exists an increasing function $h$ which satisfies

$$\lim_{t \to \infty} \frac{h(t)}{t} = \infty, \quad \sup \left\{ \int_{\Omega} h(|f|) \, d\mu : f \in \mathcal{G} \right\} < \infty.$$  

When we say $\mathcal{G}$ is bounded we mean there is some number, $M$ such that

$$\int |f| \, d\mu \leq M$$  

for all $f \in \mathcal{G}$.

7. Let $\{a_n\}, \{b_n\}$ be sequences in $[-\infty, \infty]$ and $a \in \mathbb{R}$. Show

$$\lim \inf_{n \to \infty} (a - a_n) = a - \lim \sup_{n \to \infty} a_n.$$
This was used in the proof of the Dominated convergence theorem. Also show
\[ \limsup_{n \to \infty} (-a_n) = -\liminf_{n \to \infty} (a_n) \]
\[ \limsup_{n \to \infty} (a_n + b_n) \leq \limsup_{n \to \infty} a_n + \limsup_{n \to \infty} b_n \]
provided no sum is of the form \( \infty - \infty \). Also show strict inequality can hold in the inequality. State and prove corresponding statements for \( \lim \inf \).

8. Let \( (\Omega, \mathcal{F}, \mu) \) be a measure space and suppose \( f, g : \Omega \to (-\infty, \infty] \) are measurable. Prove the sets
\[ \{ \omega : f(\omega) < g(\omega) \} \text{ and } \{ \omega : f(\omega) = g(\omega) \} \]
are measurable. **Hint:** The easy way to do this is to write
\[ \{ \omega : f(\omega) < g(\omega) \} = \bigcup_{r \in \mathbb{Q}} [f < r] \cap [g > r]. \]
Note that \( l(x, y) = x - y \) is not continuous on \( (-\infty, \infty] \) so the obvious idea doesn’t work.

9. Let \( \{ f_n \} \) be a sequence of real or complex valued measurable functions. Let
\[ S = \{ \omega : \{ f_n(\omega) \} \text{ converges} \}. \]
Show \( S \) is measurable. **Hint:** You might try to exhibit the set where \( f_n \) converges in terms of countable unions and intersections using the definition of a Cauchy sequence.

10. Let \( (\Omega, \mathcal{S}, \mu) \) be a measure space and let \( f \) be a nonnegative measurable function defined on \( \Omega \). Also let \( \phi : [0, \infty) \to [0, \infty) \) be strictly increasing and have a continuous derivative and \( \phi(0) = 0 \). Suppose \( f \) is bounded and that \( 0 \leq \phi(f(\omega)) \leq M \) for some number, \( M \). Show that
\[ \int_{\Omega} \phi(f) \, d\mu = \int_0^\infty \phi'(s) \mu([s < f]) \, ds, \]
where the integral on the right is the ordinary improper Riemann integral. **Hint:** First note that \( s \to \phi'(s) \mu([s < f]) \) is Riemann integrable because \( \phi' \) is continuous and \( s \to \mu([s < f]) \) is a nonincreasing function, hence Riemann integrable. From the second description of the Lebesgue integral and the assumption that \( \phi(f(\omega)) \leq M \), argue that for \( \lfloor M/h \rfloor \) the greatest integer less than \( M/h \),
\[ \int_{\Omega} \phi(f) \, d\mu = \sup_{h > 0} \sum_{i = 1}^{\lfloor M/h \rfloor} h \mu([ih < \phi(f)]) \]
\[ = \sup_{h > 0} \sum_{i = 1}^{\lfloor M/h \rfloor} h \mu([\phi^{-1}(ih) < f]) \]
\[ = \sup_{h > 0} \sum_{i = 1}^{\lfloor M/h \rfloor} \frac{h \Delta_i}{\Delta_i} \mu([\phi^{-1}(ih) < f]) \]
where \( \Delta_i = (\phi^{-1}(ih) - \phi^{-1}((i - 1)h)) \). Now use the mean value theorem to write

\[
\Delta_i = (\phi^{-1})'(t_i) h = \frac{1}{\phi'((\phi^{-1}(t_i))^h}
\]

for some \( t_i \) between \((i - 1)h\) and \( ih \). Therefore, the right side is of the form

\[
sup_h \sum_{i=1}^{\lfloor M/h \rfloor} \phi'(\phi^{-1}(t_i)) \Delta_i \mu(\lfloor \phi^{-1}(ih) < f \rfloor)
\]

where \( \phi^{-1}(t_i) \in (\phi^{-1}((i - 1)h), \phi^{-1}(ih)) \). Argue that if \( t_i \) were replaced with \( ih \), this would be a Riemann sum for the Riemann integral

\[
\int_0^{\phi^{-1}(M)} \phi'(t) \mu([t < f]) dt = \int_0^{\infty} \phi'(t) \mu([t < f]) dt.
\]

11. Let \((\Omega, \mathcal{F}, \mu)\) be a measure space and suppose \( f_n \) converges uniformly to \( f \) and that \( f_n \) is in \( L^1(\Omega) \). When can we conclude that

\[
\lim_{n \to \infty} \int f_n d\mu = \int f d\mu?
\]

12. Suppose \( u_n(t) \) is a differentiable function for \( t \in (a, b) \) and suppose that for \( t \in (a, b) \),

\[
|u_n(t)|, |u'_n(t)| < K_n
\]

where \( \sum_{n=1}^{\infty} K_n < \infty \). Show

\[
(\sum_{n=1}^{\infty} u_n(t))' = \sum_{n=1}^{\infty} u'_n(t).
\]

**Hint:** This is an exercise in the use of the dominated convergence theorem and the mean value theorem.

13. Show that \( \{\sum_{n=1}^{\infty} 2^{-n} \mu([i2^{-n} < f])\} \) for \( f \) a nonnegative measurable function is an increasing sequence. Could we define \( \int f d\mu = \lim_{n \to \infty} \sum_{n=1}^{\infty} 2^{-n} \mu([i2^{-n} < f]) \) and if we did, would it be equivalent to the above definitions of the Lebesgue integral?

14. Suppose \( \{f_n\} \) is a sequence of nonnegative measurable functions defined on a measure space, \((\Omega, \mathcal{S}, \mu)\). Show that

\[
\int \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int f_k d\mu.
\]

**Hint:** Use the monotone convergence theorem along with the fact the integral is linear.
The Construction Of Measures

2.1 Outer Measures

We have impressive theorems about measure spaces and the abstract Lebesgue integral. Now we need to develop interesting examples. In this chapter, we discuss the method of outer measures due to Caratheodory (1918). This approach shows how to obtain measure spaces starting with an outer measure. This will then be used to construct measures determined by positive linear functionals.

**Definition 2.1.1** Let $\Omega$ be a nonempty set and let $\mu : \mathcal{P}(\Omega) \to [0, \infty]$ satisfy

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

If $A \subseteq B$, then $\mu(A) \leq \mu(B)$,

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu(E_i).$$

Such a function is called an outer measure. For $E \subseteq \Omega$, we say $E$ is $\mu$ measurable if for all $S \subseteq \Omega$,

$$\mu(S) = \mu(S \setminus E) + \mu(S \cap E). \quad (2.1)$$

To help in remembering 2.1, think of a measurable set, $E$, as a process which divides a given set into two pieces, the part in $E$ and the part not in $E$ as in 2.1. In the Bible, there are four incidents recorded in which a process of division resulted in more stuff than was originally present.\(^1\) Measurable sets are exactly those for which no such miracle occurs. You might think of the measurable sets as the nonmiraculous sets. The idea is to show that they form a $\sigma$ algebra on which the outer measure, $\mu$, is a measure. First we give a definition and a lemma.

\(^1\) Kings 17, 2 Kings 4, Matthew 14, and Matthew 15 all contain such descriptions. The stuff involved was either oil, bread, flour or fish. In mathematics such things have also been done with sets. In the book by Bruckner Bruckner and Thompson there is an interesting discussion of the Banach Tarski paradox which says it is possible to divide a ball in $\mathbb{R}^3$ into five disjoint pieces and assemble the pieces to form two disjoint balls of the same size as the first. The details can be found in: The Banach Tarski Paradox by Wagon, Cambridge University press. 1985. It is known that all such examples must involve the axiom of choice.
Definition 2.1.2 \((\mu|_S)(A) \equiv \mu(S \cap A) \) for all \(A \subseteq \Omega\). Thus \(\mu|_S\) is the name of a new outer measure, called \(\mu\) restricted to \(S\).

The next lemma indicates that we do not lose the property of measurability by considering this restricted measure.

**Lemma 2.1.3** If \(A\) is \(\mu\) measurable, then \(A\) is \(\mu|_S\) measurable.

**Proof:** Suppose \(A\) is \(\mu\) measurable. We need to show that for all \(T \subseteq \Omega\),

\[
(\mu|_S)(T) = (\mu|_S)(T \cap A) + (\mu|_S)(T \setminus A).
\]

Thus we need to show

\[
\mu(S \cap T) = \mu(T \cap A \cap S) + \mu(T \cap S \cap A^C).
\]

But we know 2.2 holds because \(A\) is \(\mu\) measurable. Apply Definition 2.1.1 to \(S \cap T\) instead of \(S\).

If \(A\) is \(\mu|_S\) measurable, it does not follow that \(A\) is \(\mu\) measurable. Indeed, if we believe in the existence of non measurable sets, we could let \(A = S\) for such a \(\mu\) non measurable set and verify that \(S\) is \(\mu|_S\) measurable.

The next theorem is the main result on outer measures. It is a very general result which applies whenever one has an outer measure on the power set of any set. This theorem will be referred to as Carathéodory’s procedure in the rest of the book.

**Theorem 2.1.4** The collection of \(\mu\) measurable sets, \(S\), forms a \(\sigma\) algebra and

If \(F_i \in S\), \(F_i \cap F_j = \emptyset\), then \(\mu(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} \mu(F_i)\). \hfill (2.3)

If \(\cdots F_n \subseteq F_{n+1} \subseteq \cdots\), then if \(F = \bigcup_{n=1}^{\infty} F_n\) and \(F_n \in S\), it follows that

\[
\mu(F) = \lim_{n \to \infty} \mu(F_n).
\]

If \(\cdots F_n \supseteq F_{n+1} \supseteq \cdots\), and if \(F = \bigcap_{n=1}^{\infty} F_n\) for \(F_n \in S\) then if \(\mu(F_1) < \infty\), we may conclude that

\[
\mu(F) = \lim_{n \to \infty} \mu(F_n).
\]

Also, \((S, \mu)\) is complete. By this we mean that if \(F \in S\) and if \(E \subseteq \Omega\) with \(\mu(E \setminus F) + \mu(F \setminus E) = 0\), then \(E \in S\).

**Proof:** First note that \(\emptyset\) and \(\Omega\) are obviously in \(S\). Now suppose \(A, B \in S\). We show \(A \setminus B \equiv A \cap B^C\) is in \(S\). Using the assumption that \(B \in S\) in the second equation below, in which \(S \cap A\) plays the role of \(S\) in the definition for \(B\) being \(\mu\) measurable,

\[
\mu(S \cap (A \cap B^C)) + \mu(S \setminus (A \cap B^C)) = \mu(S \cap A \cap B^C) + \mu(S \cap (A^C \cup B))
\]

\[
= \mu(S \cap (A^C \cup B)) + \mu(S \cap A) - \mu(S \cap A \cap B).
\]

The following picture of \(S \cap (A^C \cup B)\) may be of use.
2.1. OUTER MEASURES

From the picture, and the measurability of $A$, we see that 2.6 is no larger than

$$\leq \mu(S \cap (A^C \cup B)) = \mu(S \cap A \cap B) + \mu(S \setminus A) + \mu(S \cap A) = \mu(S \cap A) + \mu(S \setminus A) = \mu(S).$$

This has shown that if $A, B \in \mathcal{S}$, then $A \setminus B \in \mathcal{S}$. Since $\Omega \in \mathcal{S}$, this shows that $A \in \mathcal{S}$ if and only if $A^C \in \mathcal{S}$. Now if $A, B \in \mathcal{S}, A \cup B = (A^C \cap B^C)^C = (A^C \setminus B)^C \in \mathcal{S}$. By induction, if $A_1, \cdots, A_n \in \mathcal{S}$, then so is $\cup_{i=1}^n A_i$. If $A, B \in \mathcal{S}$, with $A \cap B = \emptyset$,

$$\mu(A \cup B) = \mu((A \cup B) \cap A) + \mu((A \cup B) \setminus A) = \mu(A) + \mu(B).$$

By induction, if $A_i \cap A_j = \emptyset$ and $A_i \in \mathcal{S}, \mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$.

Now let $A = \cup_{i=1}^\infty A_i$ where $A_i \cap A_j = \emptyset$ for $i \neq j$.

$$\sum_{i=1}^\infty \mu(A_i) \geq \mu(A) \geq \mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i).$$

Since this holds for all $n$, we can take the limit as $n \to \infty$ and conclude,

$$\sum_{i=1}^\infty \mu(A_i) = \mu(A)$$

which establishes 2.3. Part 2.4 follows from part 2.3 just as in the proof of Theorem 1.1.5 on Page 2.

In order to establish 2.5, let the $F_n$ be as given there. Then, since $(F_1 \setminus F_n)$ increases to $(F_1 \setminus F)$, we may use part 2.4 to conclude

$$\lim_{n \to \infty} \left( \mu(F_1) - \mu(F_n) \right) = \mu(F_1 \setminus F).$$

Now $\mu(F_1 \setminus F) + \mu(F) \geq \mu(F_1)$ and so $\mu(F_1 \setminus F) \geq \mu(F_1) - \mu(F)$. Hence

$$\lim_{n \to \infty} \left( \mu(F_1) - \mu(F_n) \right) = \mu(F_1 \setminus F) \geq \mu(F_1) - \mu(F)$$

which implies

$$\lim_{n \to \infty} \mu(F_n) \leq \mu(F).$$

But since $F \subseteq F_n$, we also have

$$\mu(F) \leq \lim_{n \to \infty} \mu(F_n)$$

and this establishes 2.5.

It remains to show $\mathcal{S}$ is closed under countable unions. We already know that if $A \in \mathcal{S}$, then $A^C \in \mathcal{S}$ and $\mathcal{S}$ is closed under finite unions. Let $A_i \in \mathcal{S}$, $A = \cup_{i=1}^\infty A_i$, $B_n = \cup_{i=1}^n A_i$. Then

$$\mu(S) = \mu(S \cap B_n) + \mu(S \setminus B_n) = (\mu(S)(B_n) + (\mu(S)(B^C_n)).$$
By Lemma 2.1.3 we know $B_n$ is $(\mu \mid S)$ measurable and so is $B_n^C$. We want to show $\mu(S) \geq \mu(S \setminus A) + \mu(S \cap A)$. If $\mu(S) = \infty$, there is nothing to prove. Assume $\mu(S) < \infty$. Then we apply Parts 2.5 and 2.4 to the outer measure, $\mu \mid S$ in 2.7 and let $n \to \infty$. Thus

$$B_n \uparrow A, \quad B_n^C \downarrow A^C$$

and this yields $\mu(S) = (\mu(S))(A) + (\mu(S))(A^C) = \mu(S \setminus A) + \mu(S \cap A)$.

Therefore $A \in S$ and this proves Parts 2.3, 2.4, and 2.5. It remains to prove the last assertion about the measure being complete.

Let $F \in S$ and let $\mu(E \setminus F) + \mu(F \setminus E) = 0$. Then

$$\mu(S) \leq \mu(S \cap E) + \mu(S \setminus E)$$

$$= \mu(S \cap E \setminus F) + \mu(S \cap E \setminus F^C) + \mu(S \cap E^C)$$

$$\geq \mu(S \cap E \setminus F) \quad \geq \mu(S \cap E \setminus F^C) \quad \geq \mu(S \cap E^C)$$

$$\leq \mu(S \cap F) + \mu(E \setminus F) + \mu(S \setminus F) + \mu(F \setminus E)$$

$$= \mu(S \cap F) + \mu(S \setminus F) = \mu(S).$$

Hence $\mu(S) = \mu(S \cap E) + \mu(S \setminus E)$ and so $E \in S$. This shows that $(S, \mu)$ is complete and proves the theorem.

We usually encounter completeness in the following form. We have $E \subseteq F \in S$ and we know $\mu(F) = 0$. Then we conclude that $E \in S$.

Where do outer measures come from? One way to obtain an outer measure is to start with a measure $\mu$, defined on a $\sigma$ algebra of sets, $S$, and use the following definition of the outer measure induced by the measure.

**Definition 2.1.5** Let $\mu$ be a measure defined on a $\sigma$ algebra of sets, $S \subseteq P(\Omega)$. Then the outer measure induced by $\mu$, denoted by $\overline{\mu}$ is defined on $P(\Omega)$ as

$$\overline{\mu}(E) = \inf \{ \mu(F) : F \in S \text{ and } F \supseteq E \}.$$  

We also say a measure space, $(S, \Omega, \mu)$ is $\sigma$ finite if there exist measurable sets, $\Omega_i$ with $\mu(\Omega_i) < \infty$ and $\Omega = \bigcup_{i=1}^{\infty} \Omega_i$.

We leave the proof of the following lemma to the reader.

**Lemma 2.1.6** If $(S, \Omega, \mu)$ is $\sigma$ finite then there exist disjoint measurable sets, $\{B_n\}$ such that $\mu(B_n) < \infty$ and $\bigcup_{n=1}^{\infty} B_n = \Omega$.

The following lemma deals with the outer measure generated by a measure which is $\sigma$ finite. It says that if the given measure is $\sigma$ finite and complete then no new measurable sets are gained by going to the induced outer measure and then considering the measurable sets in the sense of Caratheodory.

**Lemma 2.1.7** Let $(\Omega, S, \mu)$ be any measure space and let $\overline{\mu} : P(\Omega) \to [0, \infty]$ be the outer measure induced by $\mu$. Then $\overline{\mu}$ is an outer measure as claimed and if $S$ is the set of $\overline{\mu}$ measurable sets in the sense of Caratheodory, then $S \supseteq S$ and $\overline{\mu} = \mu$ on $S$. Furthermore, if $\mu$ is $\sigma$ finite and $(\Omega, S, \mu)$ is complete, then $S = S$. 

2.1. OUTER MEASURES

Proof: It is easy to see that $\mathfrak{m}$ is an outer measure. Let $E \in \mathcal{S}$. We need to show $E \in \mathcal{S}$ and $\mathfrak{m}(E) = \mu(E)$. Let $S \subseteq \Omega$. We need to show

$$\mathfrak{m}(S) \geq \mathfrak{m}(S \cap E) + \mathfrak{m}(S \setminus E).$$

(2.8)

If $\mathfrak{m}(S) = \infty$, there is nothing to prove, so assume $\mathfrak{m}(S) < \infty$. Thus there exists $T \in \mathcal{S}$, $T \supseteq S$, and

$$\begin{align*}
\mathfrak{m}(S) &> \mu(T) - \varepsilon = \mu(T \cap E) + \mu(T \setminus E) - \varepsilon \\
&\geq \mathfrak{m}(T \cap E) + \mathfrak{m}(T \setminus E) - \varepsilon \\
&\geq \mathfrak{m}(S \cap E) + \mathfrak{m}(S \setminus E) - \varepsilon.
\end{align*}$$

Since $\varepsilon$ is arbitrary, this proves 2.8 and verifies $S \subseteq \mathcal{S}$. Now if $E \in \mathcal{S}$ and $V \supseteq E$, $\mu(E) \leq \mu(V)$. Hence, taking inf, $\mu(E) \leq \mathfrak{m}(E)$. But also $\mu(E) \geq \mathfrak{m}(E)$ since $E \in \mathcal{S}$ and $E \supseteq E$. Hence

$$\mathfrak{m}(E) \leq \mu(E) \leq \mathfrak{m}(E).$$

Now suppose $(\Omega, \mathcal{S}, \mu)$ is complete and $\sigma$ finite. Thus if $E, D \in \mathcal{S}$, and $\mu(E \setminus D) = 0$, then if $D \subseteq F \subseteq E$, it follows $F \in \mathcal{S}$ because

$$F \setminus D \subseteq E \setminus D \in \mathcal{S},$$

a set of measure zero. Therefore, $F \setminus D \in \mathcal{S}$ and so

$$F = D \cup (F \setminus D) \in \mathcal{S}.$$

We know already that $\mathcal{S} \supseteq \mathcal{S}$ so let $F \in \mathcal{S}$. Using the assumption that the measure space is $\sigma$ finite, let

$$\{B_n\} \subseteq \mathcal{S}, \cup B_n = \Omega, B_n \cap B_m = \emptyset, \mu(B_n) < \infty.$$ 

Let

$$E_n \supseteq F \cap B_n, \mu(E_n) = \mathfrak{m}(F \cap B_n),$$

(2.9)

where $E_n \in \mathcal{S}$, and let

$$H_n \supseteq B_n \setminus F = B_n \cap F^c, \mu(H_n) = \mathfrak{m}(B_n \setminus F),$$

(2.10)

where $H_n \in \mathcal{S}$. The following picture may be helpful in visualizing this situation.

![Diagram of nested sets](https://via.placeholder.com/150)

Thus $H_n \supseteq B_n \cap F^c$ and so $H_n^C \subseteq B_n^C \cup F$ which implies

$$H_n^C \cap B_n \subseteq F \cap B_n.$$
We have
\[ H^C_n \cap B_n \subseteq F \cap B_n \subseteq E_n, \quad H^C_n \cap B_n, E_n \in S. \quad (2.11) \]

**Claim:** If \( A, B, D \in \mathcal{S} \) and if \( A \supseteq B \) with \( \overline{\mu}(A \setminus B) = 0 \). Then
\[ \overline{\mu}(A \cap D) = \overline{\mu}(B \cap D). \]

**Proof of claim:** This follows from the observation that \((A \cap D) \setminus (B \cap D) \subseteq A \setminus B\).

Now from 2.9 and 2.10 and this claim,
\[
\mu(E_n \setminus (H^C_n \cap B_n)) = \overline{\mu}((F \cap B_n) \setminus (H^C_n \cap B_n)) = \overline{\mu}(F \cap B_n \cap (B^C_n \cup H_n))
\]
\[ = \overline{\mu}(F \cap H_n \cap B_n) = \overline{\mu}(F \cap (B_n \cap F^C) \cap B_n) = \overline{\mu}(\emptyset) = 0. \]

Therefore, from the assumption that \((\Omega, \mathcal{S}, \mu)\) is complete and 2.11, \( F \cap B_n \in \mathcal{S} \). Therefore,
\[ F = \bigcup_{n=1}^{\infty} F \cap B_n \in \mathcal{S}. \]

This proves the lemma.

### 2.2 Positive Linear Functionals.

One of the most important theorems related to the construction of measures is the Riesz representation theorem. In order to state this theorem, we must give the following definition. In what follows, \( \Omega \) will be a topological space. Think \( \mathbb{R}^n \) is you don’t know what one of these is. This is the example of most interest anyway.

**Definition 2.2.1** Let \( \Omega \) be a topological space. We say \( f : \Omega \to \mathbb{C} \) is in \( C_c(\Omega) \) if \( f \) is continuous and
\[ \text{spt}(f) \equiv \{ x \in \Omega : f(x) \neq 0 \} \]

is a compact set. (The symbol, \( \text{spt}(f) \) is read as “support of \( f \)”.) If we write \( C_c(V) \) for \( V \) an open set, we mean that \( \text{spt}(f) \subseteq V \).

We say \( \Lambda \) is a linear functional defined on \( C_c(\Omega) \) if \( \Lambda \) is linear,
\[ \Lambda(af + bg) = a\Lambda f + b\Lambda g \]
for all \( f, g \in C_c(\Omega) \) and \( a, b \in \mathbb{C} \).

A linear functional, \( \Lambda \), is called positive if
\[ \Lambda f \geq 0 \text{ whenever } f(x) \geq 0 \text{ for all } x \in \Omega. \]

**Definition 2.2.2** We say a topological space, \( \Omega \), is \( \sigma \)-compact if \( \Omega = \bigcup_{k=1}^{\infty} \Omega_k \) where \( \Omega_k \) is a compact subset of \( \Omega \). Note that an easy way to obtain \( \sigma \)-compact for a metric space is to simply state that the closures of balls are compact. This is because if we pick a point, \( x \in \Omega \) we can say that \( \Omega = \bigcup_{k=1}^{\infty} B(x, k) \) since any \( y \in \Omega \) is at a finite distance from \( x \) and is therefore in some \( B(x, k) \).
2.2. POSITIVE LINEAR FUNCTIONALS.

In all that follows, it is assumed $\Omega$ is a $\sigma$ compact metric space. Once again, think $\mathbb{R}^n$. From the Heine Borel theorem, $B(x, r)$ is compact.

To begin with we need some technical results and notation. In all that follows, $\Omega$ will be a metric space with the property that the closure of any open ball is compact. An obvious example of such a thing is any closed subset of $\mathbb{R}^n$ or $\mathbb{R}^n$ itself and it is these cases which interest us the most. The terminology of metric spaces is used because it is convenient and contains all the necessary ideas for the proofs which follow while being general enough to include the cases just described.

**Definition 2.2.3** If $K$ is a compact subset of an open set, $V$, we say $K \prec \phi \prec V$ if

$$\phi \in C_c(V), \phi(K) = \{1\}, \phi(\Omega) \subseteq [0, 1].$$

Also for $\phi \in C_c(\Omega)$, we say $K \prec \phi$ if

$$\phi(\Omega) \subseteq [0, 1] \text{ and } \phi(K) = 1.$$ 

We say $\phi \prec V$ if

$$\phi(\Omega) \subseteq [0, 1] \text{ and } \text{spt}(\phi) \subseteq V.$$ 

The following lemma is very useful and will be used in the next big result.

**Lemma 2.2.4** Let $S$ be any nonempty subset of a metric space, $(X, d)$ and define

$$\text{dist } (x, S) \equiv \inf \{d(x, s) : s \in S\}.$$ 

Then the mapping, $x \to \text{dist } (x, S)$ satisfies

$$|\text{dist } (y, S) - \text{dist } (x, S)| \leq d(x, y)$$

and is therefore, continuous. (This function maps the metric space, $X$ to the metric space, $\mathbb{R}$.)

**Proof:** One of $\text{dist } (y, S), \text{dist } (x, S)$ is larger than or equal to the other. Assume without loss of generality that it is $\text{dist } (y, S)$. Choose $s_1 \in S$ such that

$$\text{dist } (x, S) + \varepsilon > d(x, s_1)$$

Then

$$|\text{dist } (y, S) - \text{dist } (x, S)| = \text{dist } (y, S) - \text{dist } (x, S) \leq$$

$$d(y, s_1) - d(x, s_1) + \varepsilon \leq d(x, y) + d(x, s_1) - d(x, s_1) + \varepsilon$$

$$= d(x, y) + \varepsilon.$$ 

Since $\varepsilon$ is arbitrary, this proves the lemma.

The next theorem is a very important result known as the partition of unity theorem. Before we present it, we need a simple lemma which will be used repeatedly.
Lemma 2.2.5 \textit{Let }K\textit{ be a compact subset of the open set, }V. \textit{Then there exists an open set, }W\textit{ such that }\overline{W}\textit{ is a compact set and}

\[ K \subseteq W \subseteq \overline{W} \subseteq V. \]

\textit{Also, if }K\textit{ and }V\textit{ are as just described there exists a continuous function, }\psi\textit{ such that }K \prec \psi \prec V.\]

\textbf{Proof:} For each }k \in K\textit{, let }B(k,r_k) \equiv B_k\textit{ be such that }\overline{B_k} \subseteq V. \textit{Since }K\textit{ is compact, finitely many of these balls, }B_k, \ldots, B_h\textit{ cover }K. \textit{Let }W \equiv \cup_{i=1}^h B_{k_i}. \textit{Then it follows that }\overline{W} = \cup_{i=1}^h \overline{B_{k_i}}\textit{ and satisfies the conclusion of the lemma. Now we define }\psi\textit{ as}

\[ \psi(x) = \frac{\text{dist}(x,W^C)}{\text{dist}(x,W^C) + \text{dist}(x,K)}. \]

\textit{Note the denominator is never equal to zero because if }\text{dist}(x,W^C) = 0, \textit{then }x \in W^C\textit{ and so is at a positive distance from }K\textit{ because }W\textit{ is open and }K \subseteq W. \textit{This proves the lemma. Also note that }\text{spt}(\psi) \subseteq \overline{W}.\]

\textbf{Theorem 2.2.6 (Partition of unity)} \textit{Let }K\textit{ be a compact subset of }\Omega\textit{ and suppose}

\[ K \subseteq V = \cup_{i=1}^n V_i, \ V_i \textit{ open.} \]

\textit{Then there exist }\psi_i \prec V_i\textit{ with}

\[ \sum_{i=1}^n \psi_i(x) = 1 \]

\textit{for all }x \in K.\]

\textbf{Proof:} The first thing we do is to obtain open sets, }W_i\textit{ such that }K \subseteq \cup_{i=1}^n W_i, \ W_i \subseteq V_i, \textit{and }\overline{W_i}\textit{ is compact.}

\textit{Let }K_1 = K \setminus \cup_{i=1}^n V_i. \textit{Thus }K_1\textit{ is compact and }K_1 \subseteq V_1. \textit{By the above lemma, we let}

\[ K_1 \subseteq W_1 \subseteq \overline{W_1} \subseteq V_1 \]

\textit{with }\overline{W_1}\textit{compact and }f\textit{ be such that }K_1 \prec f \prec V_1\textit{ with}

\[ W_1 \equiv \{x : f(x) \neq 0\}. \]

\textit{Thus }W_1, W_2, \ldots, W_n\textit{ covers }K\textit{ and }\overline{W_1} \subseteq V_1. \textit{Suppose we have found }W_1, \ldots, W_r\textit{ such that each }\overline{W_r}\textit{ is compact, }\overline{W_i} \subseteq V_i\textit{ for each }i, \textit{and }W_1, \ldots, W_r, V_{r+1}, \ldots, V_n\textit{ covers }K. \textit{Then let}

\[ K_{r+1} \equiv K \setminus (\cup_{i=r+1}^n V_i) \cup (\cup_{j=1}^r W_j). \]

\textit{It follows }K_{r+1}\textit{ is compact and }K_{r+1} \subseteq V_r. \textit{Using the above lemma again, there exists }W_{r+1}\textit{ open such that}

\[ K_{r+1} \subseteq W_{r+1} \subseteq \overline{W_{r+1}} \subseteq V_{r+1} \]

\textit{and }\overline{W_{r+1}}\textit{ is compact. Continuing this way we obtain the open sets, }W_i\textit{ which were desired. Now from the above lemma, there exist }U_i\textit{ open sets}
such that $\overline{W}_i \subseteq U_i \subseteq \overline{U}_i \subseteq V_i$, $\overline{U}_i$ compact as indicated in the following picture.

![Diagram](image)

By the lemma again, we may define $\phi_i$ and $\gamma$ such that

$$\overline{U}_i \prec \phi_i \prec V_i, \quad \bigcup_{i=1}^n \overline{W}_i \prec \gamma \prec \bigcup_{i=1}^n U_i.$$ 

Now define

$$\psi_i(x) = \begin{cases} \gamma(x)\phi_i(x)/\sum_{j=1}^n \phi_j(x) & \text{if } \sum_{j=1}^n \phi_j(x) \neq 0, \\ 0 & \text{if } \sum_{j=1}^n \phi_j(x) = 0. \end{cases}$$

If $x$ is such that $\sum_{j=1}^n \phi_j(x) = 0$, then $x \notin \bigcup_{i=1}^n \overline{U}_i$ because $\phi_i$ equals one on $\overline{U}_i$. Consequently $\gamma(y) = 0$ for all $y$ near $x$ and so $\psi_i(y) = 0$ for all $y$ near $x$. Hence $\psi_i$ is continuous at such $x$. If $\sum_{j=1}^n \phi_j(x) \neq 0$, this situation persists near $x$ because each $\phi_j$ is continuous and so $\psi_i$ is continuous at such points also. Therefore $\psi_i$ is continuous. If $x \in K$, then $\gamma(x) = 1$ and so $\sum_{j=1}^n \psi_j(x) = 1$. Clearly $0 \leq \psi_i(x) \leq 1$ and $\text{sp}(\psi_j) \subseteq V_j$. This proves the theorem.

We don’t need the following corollary at this time but it is useful later.

**Corollary 2.2.7** If $H$ is a compact subset of $V_i$, we can pick our partition of unity in such a way that $\psi_i(x) = 1$ for all $x \in H$ in addition to the conclusion of Theorem 2.2.6.

**Proof:** Keep $V_i$ the same but replace $V_j$ with $\widetilde{V}_j \equiv V_j \setminus H$. Now in the proof above, applied to this modified collection of open sets, we see that if $j \neq i, \phi_j(x) = 0$ whenever $x \in H$. Therefore, $\psi_i(x) = 1$ on $H$.

Next we consider a fundamental theorem known as Caratheodory’s criterion which gives an easy to check condition which, if satisfied by an outer measure, implies that the $\sigma$ algebra of measurable sets contains the Borel sets.

**Definition 2.2.8** For two sets, $A, B$ in a metric space, we define

$$\text{dist}(A, B) \equiv \inf \{d(x, y) : x \in A, y \in B\}.$$

**Theorem 2.2.9** Let $\mu$ be an outer measure on the subsets of $(X, d)$, a metric space. If

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

whenever $\text{dist}(A, B) > 0$, then the $\sigma$ algebra of measurable sets contains the Borel sets.

**Proof:** We only need show that closed sets are in $\mathcal{S}$, the $\sigma$-algebra of measurable sets, because then the open sets are also in $\mathcal{S}$ and consequently $\mathcal{S}$ contains the Borel sets. Let $K$ be closed and let $S$ be a subset of $\Omega$. We need to show $\mu(S) \geq \mu(S \cap K) + \mu(S \setminus K)$. Therefore, we may assume without loss of generality that $\mu(S) < \infty$. Let

$$K_n \equiv \{x : \text{dist}(x, K) \leq \frac{1}{n}\}$$
By Lemma 2.2.4, \( x \to \text{dist}(x, K) \) is continuous and so \( K_n \) is closed. By the assumption of the theorem,

\[
\mu(S) \geq \mu((S \cap K) \cup (S \setminus K_n)) = \mu(S \cap K) + \mu(S \setminus K_n)
\]

(2.12)
since \( S \cap K \) and \( S \setminus K_n \) are a positive distance apart. Now

\[
\mu(S \setminus K_n) \leq \mu(S \setminus K) \leq \mu(S \setminus K_n) + \mu((K_n \setminus K) \cap S).
\]

(2.13)
If we can show \( \lim_{n \to \infty} \mu((K_n \setminus K) \cap S) = 0 \) then the theorem will be proved because this limit along with 2.13 implies \( \lim_{n \to \infty} \mu(S \setminus K_n) = \mu(S \setminus K) \) and then taking a limit in 2.12 we obtain \( \mu(S) \geq \mu(S \cap K) + \mu(S \setminus K) \) as desired. Therefore, we have reduced the proof to establishing this limit.

Since \( K \) is closed, a point, \( x \notin K \) must be at a positive distance from \( K \) and so

\[
K_n \setminus K = \bigcup_{k=n}^{\infty} K_k \setminus K_{k+1}.
\]

Therefore

\[
\mu(S \cap (K_n \setminus K)) \leq \sum_{k=n}^{\infty} \mu(S \cap (K_k \setminus K_{k+1})).
\]

(2.14)
If we can show

\[
\sum_{k=1}^{\infty} \mu(S \cap (K_k \setminus K_{k+1})) < \infty,
\]

(2.15)
then \( \mu(S \cap (K_n \setminus K)) \to 0 \) because it is dominated by the tail of a convergent series so it suffices to show 2.15.

\[
\sum_{k=1}^{M} \mu(S \cap (K_k \setminus K_{k+1})) = \sum_{k \text{ even}, \ k \leq M} \mu(S \cap (K_k \setminus K_{k+1})) + \sum_{k \text{ odd}, \ k \leq M} \mu(S \cap (K_k \setminus K_{k+1})).
\]

(2.16)
By the construction, the distance between any pair of sets, \( S \cap (K_k \setminus K_{k+1}) \) for different even values of \( k \) is positive and the distance between any pair of sets, \( S \cap (K_k \setminus K_{k+1}) \) for different odd values of \( k \) is positive. Therefore,

\[
\sum_{k \text{ even}, \ k \leq M} \mu(S \cap (K_k \setminus K_{k+1})) + \sum_{k \text{ odd}, \ k \leq M} \mu(S \cap (K_k \setminus K_{k+1})) \leq \mu(\bigcup_{k \text{ even}} S \cap (K_k \setminus K_{k+1})) + \mu(\bigcup_{k \text{ odd}} S \cap (K_k \setminus K_{k+1})) \leq 2\mu(S) < \infty
\]

and so for all \( M \), \( \sum_{k=1}^{M} \mu(S \cap (K_k \setminus K_{k+1})) \leq 2\mu(S) \) showing 2.15 and proving the theorem.

The following technical lemma will also prove useful in what follows.
2.2. POSITIVE LINEAR FUNCTIONALS.

Lemma 2.2.10 Suppose $\nu$ is a measure defined on a $\sigma$ algebra, $S$ of sets of $\Omega$, where $(\Omega, d)$ is a metric space having the property that $\Omega = \bigcup_{k=1}^{\infty} \Omega_k$ where $\Omega_k$ is a compact set and for all $k$, $\Omega_k \subseteq \Omega_{k+1}$. Suppose that $S$ contains the Borel sets and $\nu$ is finite on compact sets. Suppose that $\nu$ also has the property that for every $E \in S$,

$$\nu(E) = \inf \{ \nu(V) : V \supseteq E, \text{V open} \}.$$  \hspace{1cm} (2.17)

Then it follows that for all $E \in S$

$$\nu(E) = \sup \{ \nu(K) : K \subseteq E, \text{K compact} \}.$$ \hspace{1cm} (2.18)

Proof: Let $E \in S$ and let $l < \nu(E)$. By Theorem 1.1.5 on Page 2 we may choose $k$ large enough that

$$l < \nu(E \cap \Omega_k).$$

Now let $F \equiv \Omega_k \setminus E$. Thus $F \cup (E \cap \Omega_k) = \Omega_k$. By assumption, there is an open set, $V$ containing $F$ with

$$\nu(V) - \nu(F) = \nu(V \setminus F) < \nu(E \cap \Omega_k) - l.$$

We define the compact set, $K \equiv V^c \cap \Omega_k$. Then $K \subseteq E \cap \Omega_k$ and

$$E \cap \Omega_k \setminus K = E \cap \Omega_k \cap (V \cup \Omega_k^c)$$

$$= E \cap \Omega_k \cap V \subseteq \Omega_k \cap F^c \cap V \subseteq V \setminus F.$$

Therefore,

$$\nu(E \cap \Omega_k) - \nu(K) = \nu((E \cap \Omega_k) \setminus K)$$

$$\leq \nu(V \setminus F) < \nu(E \cap \Omega_k) - l$$

which implies

$$l < \nu(K).$$

This proves the lemma because $l < \nu(E)$ was arbitrary.

Definition 2.2.11 We say a measure which satisfies 2.17 for all $E$ measurable, is outer regular and a measure which satisfies 2.18 for all $E$ measurable is inner regular. A measure which satisfies both is called regular.

Thus Lemma 2.2.10 gives a condition under which outer regular implies inner regular.

With this preparation we are ready to prove the Riesz representation theorem for positive linear functionals.

Theorem 2.2.12 Let $(\Omega, d)$ be a metric space with the property that the closures of balls are compact and let $\Lambda$ be a positive linear functional on $C_c(\Omega)$. Then there exists a unique $\sigma$ algebra and measure, $\mu$, such that

$$\mu \text{ is complete, Borel, and regular,}$$ \hspace{1cm} (2.19)

$$\mu(K) < \infty \text{ for all } K \text{ compact,}$$ \hspace{1cm} (2.20)

$$\Lambda f = \int f d\mu \text{ for all } f \in C_c(\Omega).$$ \hspace{1cm} (2.21)

Such measures satisfying 2.19 and 2.20 are called Radon measures.
Proof: First we deal with the question of existence and then we will consider uniqueness. In all that follows \( V \) will denote an open set and \( K \) will denote a compact set. Define

\[
\mu (V) \equiv \sup \{ \Lambda (f) : f \preceq V \} , \mu (\emptyset) \equiv 0 ,
\]

(2.22)

and for an arbitrary set, \( T \),

\[
\mu (T) \equiv \inf \{ \mu (V) : V \supseteq T \} .
\]

We need to show first that this is well defined because there are two ways of defining \( \mu (V) \).

Lemma 2.2.13 \( \mu \) is a well defined outer measure on \( \mathcal{P} (\Omega) \).

Proof: First we consider whether \( \mu \) is well defined. To clarify the argument, denote by \( \mu_1 \) the first definition for open sets given in 2.22.

\[
\mu (V) \equiv \inf \{ \mu_1 (U) : U \supseteq V \} \leq \mu_1 (V).
\]

But also, whenever \( U \supseteq V \), \( \mu_1 (U) \geq \mu_1 (V) \) and so

\[
\mu (V) \geq \mu_1 (V).
\]

This proves that \( \mu \) is well defined. Next we verify \( \mu \) is an outer measure.

It is clear that if \( A \subseteq B \) then \( \mu (A) \leq \mu (B) \). We verify countable subadditivity for open sets. Thus let \( V = \bigcup_{i=1}^{\infty} V_i \) and let \( l < \mu (V) \). Then there exists \( f \preceq V \) such that \( \Lambda f > l \). Now \( \text{spt} (f) \) is a compact subset of \( V \) and so there exists \( m \) such that \( \{ V_i \}_{i=1}^{m} \) covers \( \text{spt} (f) \). Then, letting \( \psi_i \) be a partition of unity from Theorem 2.2.6 with \( \text{spt} (\psi_i) \subseteq V_i \), it follows that

\[
l < \Lambda (f) = \sum_{i=1}^{n} \Lambda (\psi_i f) \leq \sum_{i=1}^{\infty} \mu (V_i).
\]

Since \( l < \mu (V) \) is arbitrary, it follows that

\[
\mu (V) \leq \sum_{i=1}^{\infty} \mu (V_i).
\]

Now we must verify that for any sets, \( A_i \),

\[
\mu (\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu (A_i).
\]

It suffices to consider the case that \( \mu (A_i) < \infty \) for all \( i \) since if \( \mu (A_i) = \infty \) for some \( i \), there is nothing to prove. Let \( V_i \supseteq A_i \) and \( \mu (A_i) + \frac{\varepsilon}{2} > \mu (V_i) \). Then from countable subadditivity on open sets,

\[
\mu (\bigcup_{i=1}^{\infty} A_i) \leq \mu (\bigcup_{i=1}^{\infty} V_i)
\]

\[
\leq \sum_{i=1}^{\infty} \mu (V_i) \leq \sum_{i=1}^{\infty} \mu (A_i) + \frac{\varepsilon}{2} \leq \varepsilon + \sum_{i=1}^{\infty} \mu (A_i).
\]

Since \( \varepsilon \) is arbitrary, this proves the lemma.

We will denote by \( \mathcal{S} \) the \( \sigma \) algebra of \( \mu \) measurable sets.
2.2. POSITIVE LINEAR FUNCTIONALS.

Lemma 2.2.14 The outer measure, $\mu$ is finite on all compact sets and in fact, if $K < g$, then

$$\mu (K) \leq \Lambda (g) \quad (2.23)$$

Also $S$ contains the Borel sets so $\mu$ is a Borel measure.

**Proof:** Let $V_\alpha \equiv \{ x \in \Omega : g (x) > \alpha \}$ where $\alpha \in (0, 1)$ is arbitrary. Now let $\bar{h} < V_\alpha$. Thus $\bar{h} (x) \leq 1$ and equals zero off $V_\alpha$ while $\alpha^{-1} g (x) \geq 1$ on $V_\alpha$. Therefore,

$$\Lambda (\alpha^{-1} g) \geq \Lambda (h).$$

Since $\bar{h} < V_\alpha$ was arbitrary, this shows $\alpha^{-1} \Lambda (g) \geq \mu (V_\alpha) \geq \mu (K)$. Letting $\alpha \rightarrow 1$ yields the formula 2.23.

Next we verify that $S$ contains the Borel sets. First suppose that $V_1$ and $V_2$ are disjoint open sets with $\mu (V_1 \cup V_2) < \infty$. Let $f_1 < V_1$ be such that $\Lambda (f_1) + \varepsilon > \mu (V_1)$. Then

$$\mu (V_1 \cup V_2) \geq \Lambda (f_1 + f_2) = \Lambda (f_1) + \Lambda (f_2) \geq \mu (V_1) + \mu (V_2) - 2\varepsilon.$$

Since $\varepsilon$ is arbitrary, this shows that $\mu (V_1 \cup V_2) = \mu (V_1) + \mu (V_2)$.

Now suppose that $\text{dist} (A, B) = r > 0$ and $\mu (A \cup B) < \infty$. Let

$$\tilde{V}_1 \equiv \bigcup \left\{ B \left( a, \frac{r}{2} \right) : a \in A \right\}, \quad \tilde{V}_2 \equiv \bigcup \left\{ B \left( b, \frac{r}{2} \right) : b \in B \right\}.$$

Let $W$ be an open set containing $A \cup B$ such that $\mu (A \cup B) + \varepsilon > \mu (W)$. Now define $V_i \equiv W \cap \tilde{V}_i$ and $V \equiv V_1 \cup V_2$. Then

$$\mu (A \cup B) + \varepsilon > \mu (W) \geq \mu (V) = \mu (V_1) + \mu (V_2) \geq \mu (A) + \mu (B).$$

Since $\varepsilon$ is arbitrary, the conditions of Caratheodory’s criterion are satisfied showing that $S$ contains the Borel sets. This proves the lemma.

The measure space just described is complete because it comes from an outer measure using the Caratheodory procedure. Therefore, with Lemma 2.2.14 we have verified 2.19 and 2.20. The construction of $\mu$ shows outer regularity and the inner regularity follows from 2.20, shown in Lemma 2.2.14, and Lemma 2.2.10. It only remains to verify Condition 2.21, that the measure reproduces the functional in the desired manner and that the given measure and $\sigma$ algebra is unique.

Lemma 2.2.15 \( \int \int d\mu = \Lambda f \) for all $f \in C_c (\Omega)$.

**Proof:** It suffices to verify this for $f \in C_c (\Omega)$, $f$ real-valued. Suppose $f$ is such a function and $f (\Omega) \subset [a, b]$. Choose $t_0 < a$ and let $t_0 < t_1 < \cdots < t_n = b$, $t_i - t_{i-1} < \varepsilon$. Let

$$E_i = f^{-1} ((t_{i-1}, t_i]) \cap \text{spt} (f). \quad (2.24)$$

Note that $\bigcup_{i=1}^{n} E_i$ is a closed set and in fact

$$\bigcup_{i=1}^{n} E_i = \text{spt} (f) \quad (2.25)$$

since $\Omega = \bigcup_{i=1}^{n} f^{-1} ((t_{i-1}, t_i])$. From outer regularity and continuity of $f$, let $V_i \supseteq E_i$, $V_i$ is open and let $V_i$ satisfy

$$f (x) < t_i + \varepsilon \text{ for all } x \in V_i, \quad (2.26)$$
\[ \mu(V_i \setminus E_i) < \varepsilon / n. \]

By Theorem 2.2.6 there exists \( h_i \in C_c(\Omega) \) such that
\[ h_i \prec V_i, \quad \sum_{i=1}^{n} h_i(x) = 1 \text{ on spt}(f). \]

Now note that for each \( i \),
\[ f(x)h_i(x) \leq h_i(x)(t_i + \varepsilon). \]

(If \( x \in V_i \), this follows from 2.26. If \( x \notin V_i \) both sides equal 0.) Therefore,
\[
\begin{align*}
\Lambda f &= \Lambda(\sum_{i=1}^{n} f h_i) \\
&\leq \sum_{i=1}^{n} h_i(t_i + \varepsilon)(\Lambda(h_i)) \\
&= \sum_{i=1}^{n} (|t_0| + t_i + \varepsilon)(\Lambda(h_i)) - |t_0|\Lambda\left(\sum_{i=1}^{n} h_i\right).
\end{align*}
\]

Now note that \( |t_0| + t_i + \varepsilon \geq 0 \) and so from the definition of \( \mu \) and Lemma 2.2.14, this is no larger than
\[
\begin{align*}
\sum_{i=1}^{n} (|t_0| + t_i + \varepsilon)\mu(V_i) - |t_0|\mu(\text{spt}(f)) \\
\leq \sum_{i=1}^{n} (|t_0| + t_i + \varepsilon)(\mu(E_i) + \varepsilon / n) - |t_0|\mu(\text{spt}(f)) \\
\leq |t_0| \sum_{i=1}^{n} \mu(E_i) + |t_0|\varepsilon + \sum_{i=1}^{n} t_i \mu(E_i) + \varepsilon(|t_0| + |t|) \\
&\quad + \varepsilon \sum_{i=1}^{n} \mu(E_i) + \varepsilon^2 - |t_0|\mu(\text{spt}(f)).
\end{align*}
\]

From 2.25 and 2.24, the first and last terms cancel. Therefore this is no larger than
\[
\begin{align*}
(2|t_0| + |t| + \mu(\text{spt}(f)) + \varepsilon)\varepsilon + \sum_{i=1}^{n} t_i \mu(E_i) + \varepsilon \mu(\text{spt}(f)) \\
&\leq \int f d\mu + (2|t_0| + |t| + 2 \mu(\text{spt}(f)) + \varepsilon)\varepsilon.
\end{align*}
\]

Since \( \varepsilon > 0 \) is arbitrary,
\[
\Lambda f \leq \int f d\mu \quad (2.27)
\]

for all \( f \in C_c(\Omega) \), \( f \) real. Hence equality holds in 2.27 because \( \Lambda(-f) \leq -\int f d\mu \) so \( \Lambda(f) \geq \int f d\mu \). Thus \( \Lambda f = \int f d\mu \) for all \( f \in C_c(\Omega) \). Just apply the result for real functions to the real and imaginary parts of \( f \). This proves the Lemma.

Now that we have shown that \( \mu \) satisfies the conditions of the Riesz representation theorem, we show that \( \mu \) is the only measure that does so.
Lemma 2.2.16 The measure and σ algebra of Theorem 2.2.12 are unique.

Proof: If \((\mu_1, S_1)\) and \((\mu_2, S_2)\) both work, let
\[
K \subseteq V, \ K \prec f \prec V.
\]
Then
\[
\mu_1(K) \leq \int f d\mu_1 = \lambda f = \int f d\mu_2 \leq \mu_2(V).
\]
Thus \(\mu_1(K) \leq \mu_2(K)\) because of the outer regularity of \(\mu_2\). Similarly, \(\mu_1(K) \geq \mu_2(K)\) and this shows that \(\mu_1 = \mu_2\) on all compact sets. It follows from inner regularity that the two measures coincide on all open sets as well. Now let \(E \in S_1\), the σ algebra associated with \(\mu_1\), and let \(E_n = E \cap \Omega_n\). By the regularity of the measures, there exist sets \(G\) and \(H\) such that \(G\) is a countable intersection of decreasing open sets and \(H\) is a countable union of increasing compact sets which satisfy
\[
G \supseteq E_n \supseteq H, \ \mu_1(G \setminus H) = 0.
\]
Since the two measures agree on all open and compact sets, it follows that \(\mu_2(G) = \mu_1(G)\) and \(\mu_2(H) = \mu_1(H)\). Therefore \(\mu_2(G \setminus H) = \mu_1(G \setminus H) = 0\).

By completeness of \(\mu_2, E_n \in S_2\), the σ algebra associated with \(\mu_2\). Thus \(E \in S_2\) since \(E = \bigcup_{n=1}^\infty E_n\), showing that \(S_1 \subseteq S_2\). Similarly \(S_2 \subseteq S_1\). Since the two σ algebras are equal and the two measures are equal on every open set, regularity of these measures shows they coincide on all measurable sets and this proves the theorem.

The following lemma is often useful.

Lemma 2.2.17 Let \((\Omega, \mathcal{F}, \mu)\) be a measure space where \(\Omega\) is a metric space having closed balls compact or more generally a topological space. Suppose \(\mu\) is a Radon measure and \(f\) is measurable with respect to \(\mathcal{F}\). Then there exists a Borel measurable function, \(g\), such that \(g = f\) a.e.

Proof: We assume without loss of generality that \(f \geq 0\). Then let \(s_n \uparrow f\) pointwise. Say
\[
s_n(\omega) = \sum_{k=1}^{P_n} c_k^n \chi_{E_k^n}(\omega)
\]
where \(E_k^n \in \mathcal{F}\). By the outer regularity of \(\mu\), there exists a Borel set, \(F_k^n \supseteq E_k^n\) such that \(\mu(F_k^n) = \mu(E_k^n)\). In fact we can take \(F_k^n\) to be a \(G_\delta\) set if we want. Let
\[
t_n(\omega) = \sum_{k=1}^{P_n} c_k^n \chi_{F_k^n}(\omega).
\]
Then \(t_n\) is Borel measurable and \(t_n(\omega) = s_n(\omega)\) for all \(\omega \notin N_n\) where \(N_n \in \mathcal{F}\) is a set of measure zero. Now let \(N = \bigcup_{n=1}^\infty N_n\). Then \(N\) is a set of measure zero and if \(\omega \notin N\), then \(t_n(\omega) \to f(\omega)\). Let \(N' \supseteq N\) where \(N'\) is a Borel set and \(\mu(N') = 0\). Then \(t_n \chi_{N_n'}\) converges pointwise to a Borel measurable function, \(g\), and we see that \(g(\omega) = f(\omega)\) for all \(\omega \notin N'\). Therefore, \(g = f\) a.e. and this proves the lemma.
2.3 Product Measures

Let \((X, \mathcal{S}, \mu)\) and \((Y, \mathcal{T}, \nu)\) be two complete measure spaces. In this section we consider the problem of defining a product measure, \(\mu \times \nu\) which is defined on a \(\sigma\) algebra of sets of \(X \times Y\) such that \((\mu \times \nu)(E \times F) = \mu(E) \nu(F)\) whenever \(E \in \mathcal{S}\) and \(F \in \mathcal{T}\). I found the following approach to product measures in [15] and they say they got it from [17].

**Definition 2.3.1** Let \(R\) denote the set of countable unions of sets of the form \(A \times B\) where \(A \in \mathcal{S}\) and \(B \in \mathcal{T}\) and also let

\[
\rho(A \times B) = \mu(A) \nu(B) \tag{2.28}
\]

For \(S \subseteq X \times Y\), define

\[
(\mu \times \nu)(S) \equiv \inf \left\{ \sum_{i=1}^{\infty} \rho(R_i) : S \subseteq \bigcup_{i=1}^{\infty} R_i, R_i \in R \right\}. \tag{2.29}
\]

More generally, define

\[
\rho(E) \equiv \iint \chi_E(x,y) \, d\mu \, d\nu \tag{2.30}
\]

whenever \(E\) is such that

\[
x \rightarrow \chi_E(x,y) \text{ is } \mu \text{ measurable for all } y \tag{2.31}
\]

and

\[
y \rightarrow \int \chi_E(x,y) \, d\mu \text{ is } \nu \text{ measurable.} \tag{2.32}
\]

Note that if \(E = A \times B\) as above, then

\[
\int \int \chi_{A \times B}(x,y) \, d\mu \, d\nu = \int \int \chi_A(x) \, \chi_B(y) \, d\mu \, d\nu = \mu(A) \nu(B)
\]

**Lemma 2.3.2** \(\mu \times \nu\) is an outer measure on \(X \times Y\).

**Proof:** We need to verify that if \(S \subseteq T\), then

\[
(\mu \times \nu)(S) \subseteq (\mu \times \nu)(T), \tag{2.33}
\]

\[
(\mu \times \nu)(\bigcup_{i=1}^{\infty} S_i) \leq \sum_{i=1}^{\infty} (\mu \times \nu)(S_i). \tag{2.34}
\]

To do this, note that 2.33 is obvious. To verify 2.34, note that it is obvious if \((\mu \times \nu)(S_i) = \infty\) for any \(i\). Therefore, assume \((\mu \times \nu)(S_i) < \infty\). Then letting \(\varepsilon > 0\) be given, there exist

\[
\bigcup_{i=1}^{\infty} R_j^i \in R, \text{ and } (\mu \times \nu)(S_i) + \frac{\varepsilon}{2^i} > \sum_{j=1}^{\infty} \rho(R_j^i). \tag{2.35}
\]

Then \(\{R_j^i\}\) is countable and \(\bigcup_{i=1}^{\infty} S_i \subseteq \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} R_j^i\). Therefore,

\[
(\mu \times \nu)(\bigcup_{i=1}^{\infty} S_i) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho(R_j^i) \leq \sum_{i=1}^{\infty} \left( (\mu \times \nu)(S_i) + \frac{\varepsilon}{2^i} \right) = \sum_{i=1}^{\infty} (\mu \times \nu)(S_i) + \varepsilon.
\]
Since $\varepsilon > 0$ is arbitrary, this shows

$$(\mu \times \nu)(\bigcup_{i=1}^{\infty} S_i) \leq \sum_{i=1}^{\infty} (\mu \times \nu)(S_i)$$

and this proves the lemma.

By Carathéodory’s procedure, it follows there is a $\sigma$ algebra of subsets of $X \times Y$, denoted here by $\mathcal{S} \times \mathcal{T}$ such that $(\mu \times \nu)$ is a complete measure on this $\sigma$ algebra. The following lemma will be a fundamental tool in all that follows.

**Lemma 2.3.3** If $Q = \bigcup_{i=1}^{\infty} A_i \times B_i \in \mathcal{R}$, then there exist disjoint sets, of the form $A'_i \times B'_i$ such that $Q = \bigcup_{i=1}^{\infty} A'_i \times B'_i$ and $A'_i \in \mathcal{S}$ while $B'_i \in \mathcal{T}$. Also, the intersection of finitely many sets of $\mathcal{R}$ is a set of $\mathcal{R}$.

**Proof:** From the above picture, you see that if you have two sets, $A \times B$, and $C \times D$, then

$$(C \times D) \setminus (A \times B) = C \times (D \setminus B) \cup (C \setminus A) \times (D \cap B)$$

and these last two sets are disjoint.

Now consider

$$S \equiv C \times D \cup \bigcup_{i=1}^{n} A_i \times B_i$$

where the $A_i \times B_i$ are disjoint. Then, motivated by the above picture, you can replace $C \times D$ with

$$C \times (D \setminus B_1) \cup (C \setminus A_1) \times (D \cap B_1) \equiv C_1 \times D_1 \cup C_2 \times D_2$$

and

$$S = C_1 \times D_1 \cup C_2 \times D_2 \cup \bigcup_{i=1}^{n} A_i \times B_i.$$ 

Now consider a similar procedure involving $C_1 \times D_1$ and $A_2 \times B_2$. Thus replace $C_1 \times D_1$ with

$$C_1 \times (D_1 \setminus B_2) \cup (C_1 \setminus A_2) \times (D_1 \cap B_2) \equiv C_{11} \times D_{11} \cup C_{12} \times D_{12}$$

and $C_2 \times D_2$ with

$$C_2 \times (D_2 \setminus B_2) \cup (C_2 \setminus A_2) \times (D_2 \cap B_2) \equiv C_{21} \times D_{21} \cup C_{22} \times D_{22}.$$ 

Making this replacement, the union of all the sets still equals $S$ but if there are any nonempty intersections, they are between some of the above sets and $A_k \times B_k$ for some $k \geq 3$. Therefore, replace each of the above four sets,
with two sets in such a way that the resulting list of sets still has union equal to \( S \) but the only possible intersections are with sets \( A_k \times B_k \) for \( k \geq 4 \). Continuing this way, we eventually obtain
\[
S = \bigcup_{i=1}^{M} C_i \times D_i \cup \bigcup_{i=1}^{n} A_i \times B_i
\]
and no pair of sets has nonempty intersection. Note this process did not change \( \{ A_i \times B_i \}_{i=1}^{n} \).

Now let
\[
\bigcup_{i=1}^{\infty} A_i \times B_i
\]
be given. Then using the above process, there exists \( A'_i \times B'_i \) for \( i = 1, \ldots, M_1 \) such that the \( A'_i \times B'_i \) are disjoint and their union equals \( A_1 \times B_1 \cup A_2 \times B_2 \).

In fact we can take \( M_2 = 3 \) but this is not important. Now use the above procedure to obtain disjoint sets \( A'_i \times B'_i \) for \( i = 1, \ldots, M_2, \ldots, M_3 \) such there is no change in the \( A'_i \times B'_i \) for \( i \leq M_2 \) and
\[
\bigcup_{i=1}^{M_2} A'_i \times B'_i = \bigcup_{i=1}^{3} A_i \times B_i.
\]
Continue using the above procedure to generate the desired infinite list of disjoint sets recursively.

Finally
\[
\left( \bigcup_{i=1}^{\infty} A_i \times B_i \right) \cap \left( \bigcup_{j=1}^{\infty} C_j \times D_j \right) = \bigcup_{i,j} (A_i \cap C_j) \times (B_i \cap D_j)
\]
which is an element of \( \mathcal{R} \).

This proves the lemma.

**Lemma 2.3.4** For \( \rho \) defined in 2.30, 2.31 and 2.32 hold for any element of \( \mathcal{R} \). Now let \( \mathcal{R}_1 \) be defined as the set of all countable intersections of elements of \( \mathcal{R} \). For any \( S \subseteq X \times Y \),
\[
\overline{\mu \times \nu}(S) = \inf \{ \rho(P) : P \in \mathcal{R}, P \supseteq S \}
\]
and there exists \( R \in \mathcal{R}_1 \) for which 2.31 and 2.32 hold and
\[
\overline{\mu \times \nu}(S) = \rho(R). \tag{2.35}
\]
Every element of \( \mathcal{R}_1 \) is \( \overline{\mu \times \nu} \) measurable.

**Proof:** Let \( P \equiv \bigcup_{i=1}^{\infty} A_i \times B_i \in \mathcal{R} \). Then by Lemma 2.3.2, there exists \( \{ A'_i \times B'_i \}_{i=1}^{\infty} \) such that the sets are disjoint and \( \bigcup_{i=1}^{\infty} A'_i \times B'_i = P \). Therefore, since the sets are disjoint,
\[
\mathcal{X}_P(x,y) = \sum_{i=1}^{\infty} \mathcal{X}_{A'_i \times B'_i}(x,y) = \sum_{i=1}^{\infty} \mathcal{X}_{A'_i}(x) \mathcal{X}_{B'_i}(y).
\]
It follows \( x \to \mathcal{X}_P(x,y) \) is measurable. Now by the monotone convergence theorem,
\[
\int \mathcal{X}_P(x,y) \, d\mu = \int \sum_{i=1}^{\infty} \mathcal{X}_{A'_i}(x) \mathcal{X}_{B'_i}(y) \, d\mu
\]
\[
= \sum_{i=1}^{\infty} \mathcal{X}_{B'_i}(y) \int \mathcal{X}_{A'_i}(x) \, d\mu
\]
\[
= \sum_{i=1}^{\infty} \mathcal{X}_{B'_i}(y) \mu(A'_i). \]
2.3. PRODUCT MEASURES

It follows \( y \to \int X_P(x,y) \, d\mu \) is measurable and so by the monotone convergence theorem again,

\[
\int \int X_P(x,y) \, d\nu \, d\nu = \int \sum_{i=1}^{\infty} X_{P_i}(y) \mu(A_i) \, d\nu
= \sum_{i=1}^{\infty} \int X_{P_i}(y) \mu(A_i) \, d\nu
= \sum_{i=1}^{\infty} \nu(B_i) \mu(A_i).
\]  \hfill (2.36)

This shows 2.3.1 and 2.3.2 hold for \( R \in \mathcal{R} \) and also establishes 2.3.6.

Now let \( S \subseteq X \times Y \). If \( (\mu \times \nu)(S) = \infty \), let \( R = X \times Y \) and it follows \( \rho(X \times Y) = \infty = (\mu \times \nu)(S) \). Assume then that \( (\mu \times \nu)(S) < \infty \). Then there exists \( P = \bigcup_{i=1}^{\infty} A_i \times B_i \in \mathcal{R} \) such that

\[
(\mu \times \nu)(S) + \varepsilon > \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i)
= \sum_{i=1}^{\infty} \int \int X_{A_i \times B_i}(x,y) \, d\mu \, d\nu
= \int \int \sum_{i=1}^{\infty} X_{A_i \times B_i}(x,y) \, d\mu \, d\nu
\geq \int \int X_P(x,y) \, d\mu \, d\nu
= \int \int \sum_{i=1}^{\infty} X_{A_i}(x) X_{B_i}(y) \, d\mu \, d\nu
= \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i) = \rho(P) \geq (\mu \times \nu)(S).
\]

In the above, we used the monotone convergence theorem. Then we used Lemma 2.3.3 to get disjoint sets, \( \{ A'_i \times B'_i \} \) such that \( \bigcup_{i=1}^{\infty} A'_i \times B'_i = P \) and the monotone convergence theorem again. Therefore, there exists \( P_n \in \mathcal{R} \) such that \( P_n \supseteq S \) and

\[
(\mu \times \nu)(S) \leq \rho(P_n) < (\mu \times \nu)(S) + 1/n. \quad \hfill (2.37)
\]

This shows

\[
(\mu \times \nu)(S) = \inf \{ \rho(P) : P \in \mathcal{R}, P \supseteq S \}.
\]

Letting \( Q_n = \bigcap_{i=1}^{n} P_i \), it follows from Lemma 2.3.3 that \( Q_n \in \mathcal{R} \) so 2.3.7 holds for \( Q_n \) replacing \( P_n \). Also, \( Q_n \supseteq Q_{n+1} \), \( X_{Q_1} \leq X_{P_1} \), and

\[
\rho(Q_1) = \rho(P_1) = \int \int X_{P_1}(x,y) \, d\mu \, d\nu < \infty. \quad \hfill (2.38)
\]

Therefore, there exists a set of \( \nu \) measure 0, \( N \), such that if \( y \notin N \), then

\[
\int X_{P_1}(x,y) \, d\mu < \infty.
\]

Let

\[
P = \bigcap_{i=1}^{\infty} Q_i \supseteq S
\]
Then for all \((x, y) \in X \times Y\),
\[
\mathcal{X}_{Nc} (y) \mathcal{X}_P (x, y) = \lim_{n \to \infty} \mathcal{X}_{Nc} (y) \mathcal{X}_{Q_n} (x, y)
\]
and so, by the dominated convergence theorem,
\[
\int \mathcal{X}_{Nc} (y) \mathcal{X}_P (x, y) d\mu = \lim_{n \to \infty} \int \mathcal{X}_{Nc} (y) \mathcal{X}_{Q_n} (x, y) d\mu.
\]
Now since \(Q_n\) satisfies 2.31 and 2.32, \(y \to \int \mathcal{X}_{Nc} (y) \mathcal{X}_{Q_n} (x, y) d\mu\) is \(\nu\) measurable and
\[
\int \mathcal{X}_{Nc} (y) \mathcal{X}_{Q_n} (x, y) d\mu \leq \int \mathcal{X}_P (x, y) d\mu.
\]
Therefore, by 2.38 and the dominated convergence theorem, along with the completeness of the measure, \(\nu\),
\[
\int \int \mathcal{X}_P (x, y) d\mu d\nu = \int \int \mathcal{X}_{Nc} (y) \mathcal{X}_P (x, y) d\mu d\nu
\]
\[
= \lim_{n \to \infty} \int \int \mathcal{X}_{Nc} (y) \mathcal{X}_{Q_n} (x, y) d\mu d\nu
\]
\[
= \lim_{n \to \infty} \int \mathcal{X}_{Q_n} (x, y) d\mu d\nu
\]
\[
\in \left\{(\mu \times \nu) (S), \frac{\mu (S) - \nu (S)}{n}\right\}
\]
for all \(n\). Therefore,
\[
\rho (P) \equiv \int \int \mathcal{X}_P (x, y) d\mu d\nu = (\mu \times \nu) (S).
\]

Now it is necessary to verify that sets of \(\mathcal{R}_1\) are \(\mu \times \nu\) measurable. It suffices to verify that for \(A \in \mathcal{S}\) and \(B \in \mathcal{T}\), \(A \times B\) is \(\mu \times \nu\) measurable because sets of \(\mathcal{R}_1\) are just countable intersections of countable unions of these sets.

Let \(S \subseteq X \times Y\). We need to verify that
\[
(\mu \times \nu) (S) \geq (\mu \times \nu) (S \cap (A \times B)) + (\mu \times \nu) (S \setminus (A \times B)).
\]
To do so, we use the following claim.

**Claim:** Let \(P, A \times B \in \mathcal{R}\). Then
\[
\rho (P \cap (A \times B)) + \rho (P \setminus (A \times B)) = \rho (P).
\]

**Proof of the claim:** From Lemma 2.3.3, \(P = \bigcup_{i=1}^\infty A'_i \times B'_i\) where the \(A'_i \times B'_i\) are disjoint. Therefore,
\[
P \cap (A \times B) = \bigcup_{i=1}^\infty (A \cap A'_i) \times (B \cap B'_i)
\]
while
\[
P \setminus (A \times B) = \bigcup_{i=1}^\infty (A'_i \setminus A) \times B'_i \cup \bigcup_{i=1}^\infty (A \cap A'_i) \times (B'_i \setminus B).
\]
Since all of the sets in the above unions are disjoint,
\[
\rho (P \cap (A \times B)) + \rho (P \setminus (A \times B)) =
\]
\[ \int \int \sum_{i=1}^{\infty} \chi_{(A_i \cap A_i')} (x) \chi_{B \cap B_i} (y) \, d\mu \, d\nu + \int \int \sum_{i=1}^{\infty} \chi_{(A_i \setminus A_i')} (x) \chi_{B \setminus B_i} (y) \, d\mu \, d\nu \]

\[ + \int \int \sum_{i=1}^{\infty} \chi_{(A_i \cap A_i')} (x) \chi_{B_i \setminus B} (y) \, d\mu \, d\nu \]

\[ = \sum_{i=1}^{\infty} \mu (A_i \cap A_i') \nu (B \cap B_i) + \mu (A_i \setminus A_i') \nu (B_i \setminus B) \]

\[ = \sum_{i=1}^{\infty} \mu (A_i \cap A_i') \nu (B_i) + \mu (A_i \setminus A_i') \nu (B_i) = \sum_{i=1}^{\infty} \mu (A_i) \nu (B_i) = \rho (P). \]

This proves the claim.

Now continuing the proof, without loss of generality, \((\mu \times \nu) (S)\) can be assumed finite. Let \(P \supseteq S\) for \(P \in \mathcal{R}\) and

\[ (\mu \times \nu) (S) + \varepsilon > \rho (P). \]

Then from the claim,

\[ (\mu \times \nu) (S) + \varepsilon > \rho (P) = \rho (P \cap (A \times B)) + \rho (P \setminus (A \times B)) \]

\[ \geq (\mu \times \nu) (S \cap (A \times B)) + (\mu \times \nu) (S \setminus (A \times B)). \]

Since \(\varepsilon > 0\) this shows \(A \times B \) is \(\mu \times \nu\) measurable as claimed.

The following theorem is the main result.

**Theorem 2.3.5** Let \(E \subseteq X \times Y\) be \(\mu \times \nu\) measurable and suppose \((\mu \times \nu) (E) < \infty\). Then there exists a set of \(\nu\) measure 0, \(N\) and for each \(y \notin N\), a set of \(\mu\) measure 0, \(N_y\), such that

\[ (\mu \times \nu) (E) = \int \int \chi_{E \cap Y} (y) \chi_{E \setminus Y} (x) \, d\mu \, d\nu = \int \int \chi_{E} (x, y) \, d\mu \, d\nu \]

where the iterated integral on the right makes sense because for \(\nu\) a.e. \(y\), it follows \(x \rightarrow \chi_{E} (x, y)\) is \(\mu\) measurable while \(y \rightarrow \int \chi_{E} (x, y) \, d\mu\) is \(\nu\) measurable. Similarly,

\[ (\mu \times \nu) (E) = \int \int \chi_{E} (x, y) \, d\nu \, d\mu. \]

**Proof:** By Lemma 2.3.4, there exists \(R \in \mathcal{R}_1\) such that

\[ \rho (R) = (\mu \times \nu) (E), \quad R \supseteq E. \]

Therefore, since \(R\) is \(\mu \times \nu\) measurable, it follows

\[ (\mu \times \nu) (R \setminus E) = 0. \]

By Lemma 2.3.4 again, there exists \(P \supseteq R \setminus E\) with \(P \in \mathcal{R}_1\) and \(\rho (P) = 0\). Thus

\[ \int \int \chi_{P} (x, y) \, d\nu \, d\mu = 0 \]
and it follows there exists a set of $\nu$ measure zero, $N$ such that if $y \notin N$, then $\int \lambda_P(x,y)\,d\mu = 0$. It follows that for $y \notin N$, there exists a set of $\mu$ measure zero, $N_y$, such that $\lambda_P(x,y) = 0$ if $x \notin N_y$. Now consider the following picture.

![Diagram](image)

If $\lambda_{NC}(y)\lambda_{NC}(x) \neq 0$, then $y \notin N$ and $x \notin N_y$ and so $\lambda_P(x,y) = 0$. Thus $(x,y) \notin P$ and this implies $(x,y) \notin R \setminus E$. Now look at the picture. We must have either $(x,y) \in E$ or $(x,y) \notin R$. Either way,

$$\lambda_{NC}(y)\lambda_{NC}(x)\lambda_R(x,y) = \lambda_{NC}(y)\lambda_{NC}(x)\lambda_E(x,y)$$

and so, since $\mu$ and $\nu$ are complete,

$$\int \int \lambda_E(x,y)\,d\mu\,d\nu = \int \int \lambda_{NC}(y)\lambda_{NC}(x)\lambda_E(x,y)\,d\mu\,d\nu$$

$$= \int \int \lambda_{NC}(y)\lambda_{NC}(x)\lambda_R(x,y)\,d\mu\,d\nu$$

$$= \int \int \lambda_R(x,y)\,d\mu\,d\nu = \rho(R) = (\mu \times \nu)(E) \cdot$$

In all the above we could have written $d\nu d\mu$ instead of $d\mu d\nu$ and obtained the same result. This proves the theorem.

Now let $f : X \times Y \to [0, \infty]$ be $\mu \times \nu$ measurable and

$$\int f d(\mu \times \nu) < \infty. \quad (2.39)$$

Let $s(x,y) \equiv \sum_{i=1}^m c_i \lambda_{E_i}(x,y)$ be a nonnegative simple function with $c_i$ being the nonzero values of $s$ and suppose

$$0 \leq s \leq f.$$

Then from the above theorem,

$$\int sd(\mu \times \nu) = \int \int sd\mu\,d\nu$$

$$= \int \int \lambda_{NC}(y)\lambda_{NC}(x)\,s(x,y)\,d\mu\,d\nu$$

where $N$ is a finite union of sets of measure zero corresponding to the $E_i$ and $N_y$ is also such a finite union. This follows because $2.39$ implies $(\mu \times \nu)(E_i) < \infty$. Now let $s_n \uparrow f$ where $s_n$ is a nonnegative simple function and

$$\int s_n d(\mu \times \nu) = \int \int \lambda_{NC}(y)\lambda_{NC}(x)\,s_n(x,y)\,d\mu\,d\nu.$$
2.3. PRODUCT MEASURES

Then let \( N \equiv \bigcup_{n=1}^{\infty} N_n \), and \( N_y \equiv \bigcup_{n=1}^{\infty} N_{y_n} \). It follows that \( N \) is a set of \( \nu \) measure zero while for each \( y \notin N, N_y \) is a set of \( \mu \) measure zero. Thus
\[
\int s_n d(\bar{\mu} \times \nu) = \int \int X_{N^c} (y) X_{N_y^c} (x) s_n (x, y) d\mu d\nu
\]
and letting \( n \to \infty \), the monotone convergence theorem implies
\[
\int f d(\bar{\mu} \times \nu) = \int \int X_{N^c} (y) X_{N_y^c} (x) f (x, y) d\mu d\nu
= \int \int f (x, y) d\mu d\nu
\]
because of completeness of the measures, \( \mu \) and \( \nu \). This proves Fubini’s theorem.

**Theorem 2.3.6 (Fubini)** Let \((X, S, \mu)\) and \((Y, T, \nu)\) be complete measure spaces and let
\[
(\bar{\mu} \times \nu) (E) \equiv \inf \left\{ \sum_{i=1}^{\infty} \mu (A_i) \nu (B_i) : E \subset \bigcup_{i=1}^{\infty} A_i \times B_i \right\}
\]
where \(A_i \in S\) and \(B_i \in T\). Then \(\bar{\mu} \times \nu\) is an outer measure. If \(f \geq 0\) is a \(\bar{\mu} \times \nu\) measurable function satisfying
\[
\int f d(\bar{\mu} \times \nu) < \infty,
\]
then
\[
\int f d(\bar{\mu} \times \nu) = \int \int f d\mu d\nu,
\]
where the iterated integral on the right makes sense because for \(\nu\) a.e. \(y, x \to f(x, y)\) is \(\mu\) measurable and \(y \to \int f (x, y) d\mu\) is \(\nu\) measurable. Similarly,
\[
\int f d(\bar{\mu} \times \nu) = \int \int f d\nu d\mu.
\]

**Corollary 2.3.7** If \(f \in L^1 (X \times Y)\), then
\[
\int f d(\bar{\mu} \times \nu) = \int \int f (x, y) d\nu d\mu = \int \int f (x, y) d\mu d\nu.
\]

If \(\mu\) and \(\nu\) are \(\sigma\) finite, then if \(|f|\) is \(\bar{\mu} \times \nu\) measurable and either \(\int |f| d\nu d\mu < \infty\) or \(\int |f| d\mu d\nu < \infty\), then \(\int |f| d(\bar{\mu} \times \nu) < \infty\) so if \(f\) has complex values, \(f \in L^1 (X \times Y)\).

**Proof:** Without loss of generality, it can be assumed that \(f\) has real values. Then
\[
f = \frac{|f| + f - (|f| - f)}{2}
\]
and both \(f^+ \equiv \frac{|f| + f}{2}\) and \(f^- \equiv \frac{|f| - f}{2}\) are nonnegative and are less than \(|f|\). Therefore, \(\int g d(\bar{\mu} \times \nu) < \infty\) for \(g = f^+\) and \(g = f^-\) so the above theorem applies and
\[
\int f d(\bar{\mu} \times \nu) = \int f^+ d(\bar{\mu} \times \nu) - \int f^- d(\bar{\mu} \times \nu)
= \int \int f^+ d\mu d\nu - \int \int f^- d\mu d\nu
= \int \int f d\mu d\nu.
\]
It remains to verify the last claim. Suppose \( s \) is a simple function,
\[
s(x, y) = \sum_{i=1}^{m} c_i \chi_{E_i} \leq |f|(x, y)
\]
where the \( c_i \) are the nonzero values of \( s \).Then
\[
s \chi_{R_n} \leq |f| \chi_{R_n}
\]
where \( R_n \equiv X_n \times Y_n \) where \( X_n \uparrow X \) and \( Y_n \uparrow Y \) with \( \mu(X_n) < \infty \) and \( \nu(Y_n) < \infty \). It follows, since the nonzero values of \( s \chi_{R_n} \) are achieved on sets of finite measure,
\[
\int s \chi_{R_n} d(\mu \times \nu) = \int \int s \chi_{R_n} d\mu d\nu.
\]
Letting \( n \to \infty \) and applying the monotone convergence theorem, this yields
\[
\int s d(\mu \times \nu) = \int \int s d\mu d\nu. \tag{2.40}
\]
Now let \( s_n \uparrow |f| \) where \( s_n \) is a nonnegative simple function. From 2.40,
\[
\int s_n d(\mu \times \nu) = \int \int s_n d\mu d\nu.
\]
Letting \( n \to \infty \) and using the monotone convergence theorem, yields
\[
\int |f| d(\mu \times \nu) = \int \int |f| d\mu d\nu < \infty
\]

### 2.4 Exercises

1. Let \( \Omega = \mathbb{N} \), the natural numbers and let \( d(p, q) = |p - q| \), the usual distance in \( \mathbb{R} \). Show that \( (\Omega, d) \) the closures of the balls are compact. Now let \( \Lambda f \equiv \sum_{i=1}^{m} f(k) \) whenever \( f \in C_c(\Omega) \). Show this is a well defined positive linear functional on the space \( C_c(\Omega) \). Describe the measure of the Riesz representation theorem which results from this positive linear functional. What if \( \Lambda(f) = f(1) \)? What measure would result from this functional? Which functions are measurable?

2. Verify that \( \overline{\mu} \) defined in Lemma 2.1.7 is an outer measure.

3. Let \( F : \mathbb{R} \to \mathbb{R} \) be increasing and right continuous. Let \( \Lambda f \equiv \int f dF \) where the integral is the Riemann Stieltjes integral of \( f \). Show the measure \( \mu \) from the Riesz representation theorem satisfies
\[
\mu([a, b]) = F(b) - F(a^-), \mu((a, b]) = F(b) - F(a),
\]
\[
\mu([a, a]) = F(a) - F(a^-).
\]

4. Let \( \Omega \) be a metric space with the closed balls compact and suppose \( \mu \) is a measure defined on the Borel sets of \( \Omega \) which is finite on compact sets. Show there exists a unique Radon measure, \( \overline{\mu} \), which equals \( \mu \) on the Borel sets.
5. Random vectors are measurable functions, $\mathbf{X}$, mapping a probability space, $(\Omega, P, \mathcal{F})$ to $\mathbb{R}^n$. Thus $\mathbf{X}(\omega) \in \mathbb{R}^n$ for each $\omega \in \Omega$ and $P$ is a probability measure defined on the sets of $\mathcal{F}$, a $\sigma$ algebra of subsets of $\Omega$. For $E$ a Borel set in $\mathbb{R}^n$, define

$$\mu (E) \equiv P (\mathbf{X}^{-1}(E)) \equiv \text{probability that } \mathbf{X} \in E.$$ 

Show this is a well defined measure on the Borel sets of $\mathbb{R}^n$ and use Problem 4 to obtain a Radon measure, $\lambda_{\mathbf{X}}$ defined on a $\sigma$ algebra of sets of $\mathbb{R}^n$ including the Borel sets such that for $E$ a Borel set, $\lambda_{\mathbf{X}} (E) = \text{Probability that } (\mathbf{X} \in E)$.

6. Suppose $X$ and $Y$ are metric spaces having compact closed balls. Show

$$(X \times Y, d_{X \times Y})$$

is also a metric space which has the closures of balls compact. Here

$$d_{X \times Y} ((x_1, y_1), (x_2, y_2)) \equiv \max (d(x_1, x_2), d(y_1, y_2)).$$

Let

$$\mathcal{A} \equiv \{ E \times F : E \text{ is a Borel set in } X, F \text{ is a Borel set in } Y \}.$$ 

Show $\sigma(\mathcal{A})$, the smallest $\sigma$ algebra containing $\mathcal{A}$ contains the Borel sets. Hint: Show every open set in a metric space which has closed balls compact can be obtained as a countable union of compact sets. Next show this implies every open set can be obtained as a countable union of open sets of the form $U \times V$ where $U$ is open in $X$ and $V$ is open in $Y$.

7. Suppose $(\Omega, \mathcal{S}, \mu)$ is a measure space which may not be complete. Could you obtain a complete measure space, $(\Omega, \mathcal{S'}, \mu_1)$ by simply letting $\mathcal{S'}$ consist of all sets of the form $E$ where there exists $F \in \mathcal{S}$ such that $(F \setminus E) \cup (E \setminus F) \subseteq N$ for some $N \in \mathcal{S}$ which has measure zero and then let $\mu_1 (E) = \mu_1 (F)$?

8. If $\mu$ and $\nu$ are Radon measures defined on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, show $\mu \times \nu$ is also a radon measure on $\mathbb{R}^{n+m}$. Hint: Show the $\mu \times \nu$ measurable sets include the open sets using the observation that every open set in $\mathbb{R}^{n+m}$ is the countable union of sets of the form $U \times V$ where $U$ and $V$ are open in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Next verify outer regularity by considering $A \times B$ for $A, B$ measurable. Argue sets of $\mathcal{R}$ defined above have the property that they can be approximated in measure from above by open sets. Then verify the same is true of sets of $\mathcal{R}_1$. Finally conclude using an appropriate lemma that $\mu \times \nu$ is inner regular as well.

2.5 Chapter Notes

The idea of constructing a measure from a given outer measure is old, dating to the early part of the twentieth century and is due to Caratheodory. The treatment given here follows that given in Evans and Gariepy [15]. Caratheodory’s criterion for a measure to be Borel is then used to give a fairly short proof of the Riesz representation theorem for positive linear
functionals defined on $C_c(X)$ where $X$ is a metric space in which the balls have compact closure. More general theorems are available in Rudin [41], Hewitt and Stromberg [26], Lang [32], and Kutzler [31]. These theorems only require $X$ to be a locally compact Hausdorff space, but they do involve more intricate topological considerations. The proof of this important result given here uses the notation and some of the arguments for the more general theorem presented in Rudin [41] but it is quite a bit easier because it is less ambitious. Other ways to do Fubini’s theorem involve the notions of algebras and monotone classes. See Rudin [41], Kutzler [31], or Hewitt and Stromberg [26]. This other approach is more elegant but the results are slightly inferior to the above method which I found in Evans and Gariepy [15] and they got from Federer [17].
Bibliography


