Appendix B

Background Material: Measure Theory and integration

In this appendix we shall briefly review some of the basics of measure theory, including sigma algebras, measurable spaces, measures and the construction of the Lebesgue measure. These notes follow parts of the notes from Prof. Norris' version of the course *Probability and measure*, as well as the books *Real and complex analysis* by Rudin, *Real Analysis* by Stein and Shakarchi annd *Measure Theory and Integration* by M. Taylor.

B.1 Sigma algebras and measures

Given a set E, the basic goal of measure theory is to assign to certain subsets $A \subset E$ a value, $\mu(A)$ which represents in some appropriate sense the 'size' of A. For example, if E is finite or countable and A is any subset of E we might set $\mu(A)$ to be the number of elements in A (where $\mu(A)$ may be ∞ if A is not finite). In this case μ is defined on all of the power set 2^E. We call μ the *counting measure*.

For $E = \mathbb{R}$, it is natural to wish to define $\mu(A)$ to be the 'length' of A. This is unambiguous if A is some interval, but it turns out that we run into problems trying to define the 'length' of an arbitrary subset of $\mathbb{R}$. As a consequence, we will need to restrict our attention to a smaller collection of sets than the power set $2^{\mathbb{R}}$.

Definition B.1. *Let E be a set. A collection $\mathcal{E}$ of subsets of E is called a* σ-algebra[1] *if $\mathcal{E}$ contains $\emptyset$ and is closed under taking the complement and forming countable unions. That is if $A \in \mathcal{E}$ then*

$$A^c = \{x \in E | x \notin A\} \in \mathcal{E},$$

and if $(A_n)_{n=1}^{\infty}$ is a sequence with $A_n \in \mathcal{E}$, then

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{E}.$$

$(E, \mathcal{E})$ *is called a* measurable space.

[1] pronounced "sigma algebra"

A measure *on* $(E, \mathcal{E})$ *is a set function* $\mu : \mathcal{E} \to [0, \infty]$ *such that* $\mu(\emptyset) = 0$ *and* μ *is* countably additive. *That is for a sequence* $(A_n)_{n=1}^{\infty}$ *with* $A_n \in \mathcal{E}$ *disjoint, we have:*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

$(E, \mathcal{E}, \mu)$ *is called a* measure space.

Note that $(E, 2^E)$ is always a measurable space, since 2^E is always a σ-algebra. Suppose $(E, \mathcal{E}, \mu)$ is a measure space and that $A \in \mathcal{E}$. Then we can define a new measure space $(A, \mathcal{E}|_A, \mu|_A)$ by taking $\mathcal{E}|_A = \{B \in \mathcal{E} : B \subset A\}$ and defining $\mu|_A(B) = \mu(B)$ for all $B \in \mathcal{E}|_A$.

Exercise B.1. Let E be finite or countable and $\mathcal{E} = 2^E$.

a) Verify that if μ is the counting measure, then $(E, \mathcal{E}, \mu)$ is a measure space.

b) A mass function is a map $m : E \to [0, \infty]$. Define a set-function on $(E, \mathcal{E})$ by

$$\mu_m(A) = \sum_{x \in A} m(x).$$

Show that μ_m is a measure on $(E, \mathcal{E})$, and moreover if μ is any measure on $(E, \mathcal{E})$ then $\mu = \mu_m$ for some m.

For the examples in Exercise B.1, we can identify in a straightforward way both an appropriate σ-algebra and measure. In more general situations we may not be so lucky, so it is very helpful to be able to appeal to abstract results to construct measure spaces by starting with something simpler. We shall require the following Lemma, whose proof we defer to an exercise.

Lemma B.1. *Suppose that for each* $i \in I$*, where* I *is some (not necessarily countable) index set,* $\mathcal{E}_i$ *is a* σ*-algebra of the set* E*. Then the intersection* $\cap_{i \in I} \mathcal{E}_i$ *is a* σ*-algebra.*

With this fact in hand we can define the σ-algebra generated by a collection of sets.

Definition B.2. *If* $\mathcal{A}$ *is a collection of subsets of* E*, then the* σ-algebra generated by $\mathcal{A}$*, denoted* $\sigma(\mathcal{A})$*, is the intersection of all* σ*-algebras* $\mathcal{E}$ *on* E *such that* $\mathcal{A} \subset \mathcal{E}$*.*

Since 2^E is always a σ-algebra, and $\mathcal{A} \subset 2^E$, $\sigma(\mathcal{A})$ is always well defined. When (E, τ) is a topological space, with τ the collection of open sets, it is natural to introduce the *Borel algebra*[2] $\mathcal{B}(E) := \sigma(\tau)$. When $E = \mathbb{R}$ with it's standard topology, we often write $\mathcal{B} := \mathcal{B}(\mathbb{R})$. A measure defined on the measure space $(E, \mathcal{B}(E))$ is called a *Borel measure*. A Borel measure which is finite on compact sets is called a *Radon measure*.

Exercise B.2. a) Prove Lemma B.1.

b) Let $E = \{1, 2, 3\}$. Find $\sigma(\{1\})$, and show that $\sigma(\{1\}) \neq 2^E$.

[2]The notation $\mathcal{B}(E)$ assumes that the topology is obvious from context

Exercise B.3. a) Show that if $U \subset \mathbb{R}$ is open in the standard topology, then:

$$U = \bigcup_{n=1}^{\infty} I_n,$$

where each $I_n = (a_n, b_n)$ with $a_n < b_n$ is an open interval, and the I_n's are disjoint.

b) Show that $\mathcal{B} = \sigma(\mathcal{A})$ when $\mathcal{A}$ is given by:

i) $\mathcal{A} = \{(a,b)|a,b \in \mathbb{R}, a < b\}$, the collection of all open intervals in $\mathbb{R}$.

ii) $\mathcal{A} = \{[a,b]|a,b \in \mathbb{R}, a < b\}$, the collection of all closed intervals in $\mathbb{R}$.

iii) $\mathcal{A} = \{(a,b]|a,b \in \mathbb{R}, a < b\}$.

iv) $\mathcal{A} = \{[a,b)|a,b \in \mathbb{R}, a < b\}$.

v) $\mathcal{A} = \{(-\infty,b)|b \in \mathbb{R}\}$.

vi) $\mathcal{A} = \{(-\infty,b]|b \in \mathbb{R}\}$.

vii) $\mathcal{A} = \{(a,\infty)|a \in \mathbb{R}\}$.

viii) $\mathcal{A} = \{[a,\infty)|a \in \mathbb{R}, a\}$.

ix) $\mathcal{A} = \{(a,b)|a,b \in \mathbb{Q}, a < b\}$.

[Hint : reduce cases ii) − ix) to case i).]

We have a means of generating a σ-algebra from a smaller collection of sets, $\mathcal{A}$. We'd like to define a measure by how it acts on $\mathcal{A}$, and then 'extend' this measure to act on $\sigma(\mathcal{A})$ (or some larger σ-algebra containing $\mathcal{A}$. For this we need both an existence and a uniqueness result for the extension. We first introduce the idea of π-system and d-system and establish Dynkin's π-system Lemma, which will eventually furnish a proof of uniqueness for extensions of measures.

Definition B.3. *Let $\mathcal{A}$ be a collection of subsets of E. We say that*

i) $\mathcal{A}$ is a π-system if it contains the empty set and is closed under pairwise intersection, i.e.

- $\emptyset \in \mathcal{A}$,
- $A \cap B \in \mathcal{A}$ *for all* $A, B \in \mathcal{A}$.

ii) $\mathcal{A}$ is a d-system if it contains E and is closed under taking differences, and countable unions of increasing sets, i.e.

- $E \in \mathcal{A}$,
- $B \setminus A \in \mathcal{A}$ *for all* $A, B \in \mathcal{A}$ *with* $A \subset B$,
- $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ *for all sequences* $(A_n)_{n=1}^{\infty}$ *with* $A_n \in \mathcal{A}$ *and* $A_n \subset A_{n+1}$.

Exercise($*$). Show that if $\mathcal{A}$ is both a π-system and a d-system, then it is a σ-algebra.

Dynkin's π-system Lemma extends the previous exercise.

Lemma B.2 (Dynkin's π-system Lemma)**.** *Let $\mathcal{A}$ be a π-system. Then any d-system containing $\mathcal{A}$ also contains $\sigma(\mathcal{A})$.*

Proof. Let $\mathcal{D}$ denote the intersection of all d-systems containing $\mathcal{A}$. Then $\mathcal{D}$ is a d-system and it suffices to show that $\sigma(\mathcal{A}) \subset \mathcal{D}$. In order to do this, we show that $\mathcal{D}$ is a π-system, hence it is a σ-algebra containing $\mathcal{A}$ and thus it must contain $\sigma(\mathcal{A})$ from the definition of $\sigma(\mathcal{A})$.

We introduce

$$\mathcal{D}' = \{B \in \mathcal{D} | B \cap A \in \mathcal{D} \text{ for all } A \in \mathcal{A}\}.$$

Clearly $\mathcal{A} \subset \mathcal{D}'$ because $\mathcal{A}$ is a π-system. Next we claim that $\mathcal{D}'$ is a d-system. Clearly $E \in \mathcal{D}'$. Suppose $B_1, B_2 \in \mathcal{D}'$ with $B_1 \subset B_2$, then for $A \in \mathcal{A}$ we have:

$$(B_2 \setminus B_1) \cap A = (B_2 \cap A) \setminus (B_1 \cap A) \in \mathcal{D}$$

because $\mathcal{D}$ is a d-system and we conclude $B_2 \setminus B_1 \in \mathcal{D}'$. Now suppose $B_n \in \mathcal{D}'$ and $B_n \subset B_{n+1}$ and let $B = \bigcup_{n=1}^{\infty} B_n$. Then for any $A \in \mathcal{A}$, we have $C_n := B_n \cap A \in \mathcal{D}$, $C_n \subset C_{n+1}$ so $\bigcup_{n=1}^{\infty} C_n = B \cap A \in \mathcal{D}$ as $\mathcal{D}$ is a d-system. We deduce that $\mathcal{D}'$ is a d-system containing $\mathcal{A}$, hence $\mathcal{D}' = \mathcal{D}$ by the minimality of $\mathcal{D}$.

Now, we let

$$\mathcal{D}'' = \{B \in \mathcal{D} | B \cap A \in \mathcal{D} \text{ for all } A \in \mathcal{D}\}.$$

By the above, we have that $\mathcal{A} \subset \mathcal{D}''$, since $\mathcal{D}' = \mathcal{D}$. By the same arguments as above we can check that $\mathcal{D}''$ is a d-system, and so $\mathcal{D}'' = \mathcal{D}$ and $\mathcal{D}''$ is a π-system as required. □

B.1.1 Construction of measures

As described above, we are going to give a means of constructing a measure by specifying how it behaves on some suitable collection of sets. First we introduce some notation concerning *set functions*

Definition B.4. *Let $\mathcal{A}$ be a collection of subsets of E containing $\emptyset$. A* set function *is a function $\mu : \mathcal{A} \to [0, \infty]$ with $\mu(\emptyset) = 0$. We say that a set function μ is:*

- increasing *if*
$$\mu(A) \leqslant \mu(B), \quad \text{for all } A, B \in \mathcal{A}, \text{ with } A \subset B,$$
- additive *if, for all* ***disjoint*** *sets $A, B \in \mathcal{A}$ with $A \cup B \in \mathcal{A}$ we have:*
$$\mu(A \cup B) = \mu(A) + \mu(B),$$
- countably additive *if for all sequence of* ***disjoint*** *sets $(A_n)_{n=1}^{\infty}$ with $A_n \in \mathcal{A}$ and $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$ we have:*
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n),$$

- countably subadditive *if for all sequences* $(A_n)_{n=1}^{\infty}$ *with* $A_n \in \mathcal{A}$ *and* $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$ *we have:*
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu(A_n),$$

We shall also need to define what it means for a collection of subsets to be a ring

Definition B.5. *Let* $\mathcal{A}$ *be a collection of subsets of* E*. We say* $\mathcal{A}$ *is a* ring *on* E *if* $\emptyset \in \mathcal{A}$ *and for all* $A, B \in \mathcal{A}$*:*
$$B \setminus A \in \mathcal{A}, \quad A \cup B \in \mathcal{A}.$$
We say $\mathcal{A}$ *is an* algebra *if* $\emptyset \in \mathcal{A}$ *and for all* $A, B \in \mathcal{A}$*:*
$$A^c \in \mathcal{A}, \quad A \cup B \in \mathcal{A}.$$

Let us suppose that $\mathcal{A}$ is a ring of subsets of E, together with a countably additive set function $\mu : \mathcal{A} \to [0, \infty]$. For any set $B \subset E$, we can introduce the *outer measure*
$$\mu^*(B) := \inf \sum_{n=1}^{\infty} \mu(A_n),$$
where the infimum is taken over all sequences $(A_n)_{n=1}^{\infty}$ of sets such that $A_n \in \mathcal{A}$ and $B \subset \bigcup_{n=1}^{\infty} A_n$. If no such sequence exists we set $\mu^*(B) = \infty$. We clearly have $\mu^*(\emptyset) = 0$, so we have a set function defined on 2^E and moreover, μ^* is increasing. In general, however, μ^* will not define a measure on the measure space $(E, 2^E)$, in order for μ^* to be a measure we must restrict to a smaller σ-algebra. We say that $A \subset E$ is μ^*-measurable if, for all $B \subset E$ we have:
$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c),$$
and we denote by $\mathcal{M}$ the collection of all μ^*-measurable sets. One of the fundamental results of measure theory is:

Theorem B.3 (Carathéodory's Theorem)**.** *Suppose* $\mathcal{A}$ *is a ring of subsets of* E*, and* $\mu : \mathcal{A} \to [0, \infty]$ *is a countably additive set function. Define* $\mu^*, \mathcal{M}$ *as above. The collection* $\mathcal{M}$ *is a* σ*-algebra which contains* $\mathcal{A}$*. The set function* $\mu^* : \mathcal{M} \to [0, \infty]$ *is a measure on* $(E, \mathcal{M})$*.*

We shall establish this result through several Lemmas. First, we establish countable subadditivity of μ^*.

Lemma B.4. *The set function* $\mu^* : 2^E \to [0, \infty]$ *is countably subadditive.*

Proof. Let $B = \cup_{n=1}^{\infty} B_n$. We wish to show
$$\mu^*(B) \leqslant \sum_{n=1}^{\infty} (B_n).$$

We can easily see that if $\mu^*(B_n) = \infty$ for some n, then necessarily $\mu^*(B) = \infty$, so we can focus on the case where $\mu^*(B_n) < \infty$ for all n. Fix $\epsilon > 0$. For each n we can find a sequence of sets $(A_{n,m})_{m=1}^{\infty}$ such that $A_{n,m} \in \mathcal{A}$ with $B_n \subset \cup_{m=1}^{\infty} A_{n,m}$ and

$$\sum_{m=1}^{\infty} \mu(A_{n,m}) \leqslant \mu^*(B_n) + \epsilon 2^{-n}.$$

Now, $B \subset \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} A_{n,m}$, so we have:

$$\mu^*(B) \leqslant \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu(A_{n,m}) \leqslant \sum_{n=1}^{\infty} \mu(B_n) + \epsilon.$$

Since ϵ was arbitrary, the result follows. □

Next we show that μ^* extends μ.

Lemma B.5. *Suppose $A \in \mathcal{A}$. Then $\mu^*(A) = \mu(A)$.*

Proof. It is obvious that $\mu^*(A) \leqslant \mu(A)$, by considering the sequence $A_1 = A$, $A_n = \emptyset$ for $n > 1$, so it suffices to show $\mu^*(A) \geqslant \mu(A)$. Since μ is countably additive, it is finitely additive (take all but finitely many elements of the sequence to be the empty set). Since $\mathcal{A}$ is a ring, if $A, B \in \mathcal{A}$ with $A \subset B$, then $B \setminus A \in \mathcal{A}$. By finite additivity of μ:

$$\mu(B) = \mu\left(A \cup (B \setminus A)\right) = \mu(A) + \mu(B \setminus A) \geqslant \mu(A)$$

so μ is increasing. Suppose $(A_n)_{n=1}^{\infty}$ is a sequence with $A_n \in \mathcal{A}$. Let $B_1 = A_1$ and

$$B_n = \bigcup_{k=1}^{n} A_k \setminus \bigcup_{k=1}^{n-1} A_k$$

for $n > 1$. Then $(B_n)_{n=1}^{\infty}$ is a disjoint sequence, $B_n \subset A_n$ and moreover each $B_n \in \mathcal{A}$ since $\mathcal{A}$ is a ring. We have:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n),$$

so μ is countably subadditive.

Now, suppose $A \in \mathcal{A}$ and take any sequence $(A_n)_{n=1}^{\infty}$ with $A_n \in \mathcal{A}$ and $A \subset \cup_{n=1}^{\infty} A_n$. Note that $A \cap A_n = A \setminus ((A \cup A_n) \setminus A)$, so $A \cap A_n \in \mathcal{A}$. We deduce:

$$\mu(A) = \mu\left(\bigcup_{n=1}^{\infty} (A \cap A_n)\right) \leqslant \sum_{n=1}^{\infty} \mu(A \cap A_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n)$$

Taking the infimum over all such sequences, we conclude $\mu(A) \leqslant \mu^*(A)$ and we're done. □

Lemma B.6. *$\mathcal{M}$ contains $\mathcal{A}$.*

Proof. Suppose $A \in \mathcal{A}$ and $B \subset E$. We need to show:

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c).$$

Since $B = (B \cap A) \cup (B \cap A^c)$ and using subadditivity of μ^*, it is immediate that $\mu^*(B) \leqslant \mu^*(B \cap A) + \mu^*(B \cap A^c)$, so it suffices to show

$$\mu^*(B) \geqslant \mu^*(B \cap A) + \mu^*(B \cap A^c).$$

If $\mu^*(B) = \infty$ this is trivial, so we can focus on the case $\mu^*(B) < \infty$. Fix $\epsilon > 0$, then there exists a sequence $(A_n)_{n=1}^{\infty}$ with $A_n \in \mathcal{A}$, $B \subset \cup_{n=1}^{\infty} A_n$ and

$$\sum_{n=1}^{\infty} \mu(A_n) \leqslant \mu^*(B) + \epsilon.$$

We note that:

$$B \cap A \subset \bigcup_{n=1}^{\infty} (A_n \cap A), \qquad B \cap A^c \subset \bigcup_{n=1}^{\infty} (A_n \cap A^c).$$

Recalling that $A_n \cap A \in \mathcal{A}$ and noting that $A_n \cap A^c = (A \cup A_n) \setminus A \in \mathcal{A}$ we deduce:

$$\mu^*(B \cap A) + \mu^*(B \cap A^c) \leqslant \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \cap A^c) = \sum_{n=1}^{\infty} \mu(A_n) \leqslant \mu^*(B) + \epsilon.$$

Since ϵ was arbitrary, we're done. □

Lemma B.7. *$\mathcal{M}$ is an algebra.*

Proof. From the definition of $\mathcal{M}$ it is immediate that $E \in \mathcal{M}$ and that $A \in \mathcal{M}$ implies $A^c \in \mathcal{M}$. It remains to show that $\mathcal{M}$ is closed under pairwise union, or equivalently pairwise intersection (since $A \cup B = (A^c \cap B^c)^c$). Suppose that $A_1, A_2 \in \mathcal{M}$ and $B \subset E$. Then

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap (A_1 \cap A_2)^c \cap A_1) + \mu^*(B \cap (A_1 \cap A_2)^c \cap A_1^c) \\
&= \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap (A_1 \cap A_2)^c)
\end{aligned}$$

so $A_1 \cap A_2 \in \mathcal{M}$. □

Finally, we are ready to prove Carathéodory's theorem:

Proof of Theorem B.3. We already know that $\mathcal{M}$ is an algebra containing $\mathcal{A}$, so it suffices to show that if $(A_n)_{n=1}^{\infty}$ is a sequence of disjoint sets with $A_n \in \mathcal{M}$, and $A = \bigcup_{n=1}^{\infty} A_n$, then we have:

$$A \in \mathcal{A}, \qquad \mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

so that $\mathcal{M}$ is closed under countable unions and hence is a σ-algebra, and μ^* is a countably additive set function on $\mathcal{M}$, hence a measure. Fix any $B \subset E$. Since the A_n are disjoint, we know $A_1 \cap A_2 = \emptyset$ and $A_1 \cap A_2^c = A_1$. We deduce:

$$\begin{aligned}
\mu^*(B) &= \mu^*(B \cap A_1) + \mu^*(B \cap A_1^c) \\
&= \mu^*(B \cap A_1 \cap A_2) + \mu^*(B \cap A_1 \cap A_2^c) + \mu^*(B \cap A_1^c \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) \\
&= \mu^*(B \cap A_1) + \mu^*(B \cap A_2) + \mu^*(B \cap A_1^c \cap A_2^c) \\
&= \ldots = \sum_{k=1}^{n} \mu^*(B \cap A_k) + \mu^*(B \cap A_1^c \cap \cdots \cap A_n^c)
\end{aligned}$$

Now, since $B \cap A^c \subset B \cap A_1^c \cap \cdots \cap A_n^c$, by the fact that μ^* is increasing we know $\mu^*(B \cap A_1^c \cap \cdots \cap A_n^c) \geqslant \mu^*(B \cap A^c)$. Hence, letting $n \to \infty$ and using countable subadditivity we find:

$$\mu^*(B) \geqslant \sum_{k=1}^{\infty} \mu^*(B \cap A_k) + \mu^*(B \cap A^c) \geqslant \mu^*(B \cap A) + \mu^*(B \cap A^c) \tag{B.1}$$

The reverse inequality holds by subadditivity, and so we have

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \cap A^c)$$

and thus $A \in \mathcal{M}$. Setting $B = A$ in (B.1) we deduce:

$$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n). \qquad \square$$

Carathéodory's theorem gives a way to extend a countably additive set function defined on a ring $\mathcal{A}$ to a measure on $\sigma(\mathcal{A})$, since we can restrict the outer measure to $\sigma(\mathcal{A})$. It is often useful to know whether this extension of μ is unique. We have the following result:

Theorem B.8. *Let μ_1, μ_2 be measures on $(E, \mathcal{E})$ with $\mu_1(E) = \mu_2(E) < \infty$. Suppose that $\mu_1 = \mu_2$ on $\mathcal{A}$, where $\mathcal{A}$ is a π-system which generates $\mathcal{E}$. Then $\mu_1 = \mu_2$ on $\mathcal{E}$.*

Proof. Let $\mathcal{D} = \{A \in \mathcal{E} | \mu_1(A) = \mu_2(A)\}$ be the collection of sets on which the measures agree. By hypothesis $E \in \mathcal{D}$ and $\mathcal{A} \subset \mathcal{D}$. We shall show that $\mathcal{D}$ is a d-system, so by Dynkin's π-system Lemma we have $\mathcal{E} = \sigma(\mathcal{A}) \subset \mathcal{D}$ and we're done.

Suppose $A, B \in \mathcal{E}$ with $A \subset B$, then by additivity of the measures, we have:

$$\mu_1(A) + \mu_1(B \setminus A) = \mu_1(B) < \infty, \quad \mu_2(A) + \mu_2(B \setminus A) = \mu_2(B) < \infty,$$

so that if $A, B \in \mathcal{D}$ then $B \setminus A \in \mathcal{D}$.

Now suppose that we have a sequence $(A_n)_{n=1}^{\infty}$ with $A_n \in \mathcal{D}$ and $A_n \subset A_{n+1}$ and $A = \cup_{n=1}^{\infty} A_n$. Then setting $B_1 = A_1$ and $B_n = A_n \setminus A_{n-1}$ for $n > 1$, we can write $A = \cup_{n=1}^{\infty} B_n$, where the B_n are disjoint. Thus:

$$\mu_1(A) = \sum_{n=1}^{\infty} \mu_1(B_n) = \sum_{n=1}^{\infty} \mu_2(B_n) = \mu_2(A)$$

and hence $A \in \mathcal{D}$. Thus $\mathcal{D}$ is a d-system and so $\mathcal{E} = \mathcal{D}$ and we're done. $\square$

This result requires that E has *finite* measure. For many of the situations we're interested in this is too restrictive an assumption. We can extend the result for measures which satisfy a weaker condition.

Corollary B.9. *Let μ_1, μ_2 be measures on $(E, \mathcal{E})$. Suppose that $\mu_1 = \mu_2$ on $\mathcal{A}$, where $\mathcal{A}$ is a π-system which generates $\mathcal{E}$. Suppose also that $E = \bigcup_{i=1}^{\infty} B_i$, where $B_i \in \mathcal{A}$ and the B_i's are disjoint with $\mu_1(B_i) = \mu_2(B_i) < \infty$. Then $\mu_1 = \mu_2$ on $\mathcal{E}$.*

Proof. For each i, and for any $A \in \mathcal{E}$, define $\mu_1^i(A) = \mu_1(A \cap B_i)$, $\mu_2^i(A) = \mu_2(A \cap B_i)$. By assumption we have $\mu_1^i(E) = \mu_2^i(E) < \infty$ and moreover $\mu_1^i(A) = \mu_2^i(A)$ for all $A \in \mathcal{A}$. Thus $\mu_1^i = \mu_2^i$ on $\mathcal{E}$. Further, if $A \in \mathcal{E}$ is any measurable set, then

$$\begin{aligned}\mu_1(A) = \mu_1\left(\bigcup_{i=1}^{\infty}(B_i \cap A)\right) &= \sum_{i=1}^{\infty} \mu_1(B_i \cap A)\\ &= \sum_{i=1}^{\infty} \mu_2(B_i \cap A) = \mu_2\left(\bigcup_{i=1}^{\infty}(B_i \cap A)\right) = \mu_2(A)\end{aligned}$$

□

Completeness of measures

A useful feature of the measures obtained from Carathéodory's theorem is that they have a property known as *completeness*.

Definition B.6. *Let $(E, \mathcal{E}, \mu)$ be a measure space. We say μ is* complete *if for any $A \in \mathcal{E}$ with $\mu(A) = 0$, each subset of A also belongs to $\mathcal{E}$.*

A subset of a set of measure zero is sometimes known as a *null set*, so a complete measure is one for which all null sets are measurable.

Lemma B.10. *Suppose $(E, \mathcal{M}, \mu)$ is a measure space obtained from Carathéodory's theorem. Then it is complete.*

Proof. Let μ^* be the outer measure on E whose restriction to $\mathcal{M}$ gives μ. Suppose $N \subset A$, where $A \in \mathcal{M}$ with $\mu(A) = 0$. Since μ^* is increasing we have $\mu^*(N) \leqslant \mu(A) = 0$, so $\mu^*(N) = 0$. For any set $B \subset E$ we have:

$$\mu^*(T \cap N) + \mu^*(T \cap N^c) \leqslant \mu^*(N) + \mu^*(T) = \mu^*(T)$$

again using the increasing property of μ^*. By Lemma B.4 we know μ^* is subadditive, hence

$$\mu^*(T) \leqslant \mu^*(T \cap N) + \mu^*(T \cap N^c),$$

and thus $N \in \mathcal{M}$. □

B.1.2 Lebesgue measure

We specialise now to (arguably) the most important measure, the Lebesgue measure. This measure gives us the standard notion of volume for sets in $\mathbb{R}^n$. We first introduce the *rectangles* in $\mathbb{R}^n$.

Definition B.7. *A* rectangle *in $\mathbb{R}^n$ is a set of the form:*

$$R = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n],$$

with $a_i < b_i$ for $i = 1, \dots, n$. We define $\mathcal{A}_R$ to be the collection of finite unions of disjoint rectangles.

> **Exercise(∗).** Show that:
>
> a) The collection of rectangles is a π-system.
>
> b) $\mathcal{A}_R$ is a ring.
>
> c) $\mathcal{A}_R$ generates $\mathcal{B}(\mathbb{R}^n)$.

The main result we will establish shows:

Theorem B.11. *There exists a unique Borel measure μ on $\mathbb{R}^n$ such that, for all rectangles $R = (a_1, b_1] \times \cdots \times (a_n, b_n]$ with $a_i < b_i$ for $i = 1, \dots, n$,*

$$\mu(R) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

The measure μ is called the Lebesgue measure on $\mathbb{R}^n$.

Proof. For any $A \in \mathcal{A}_R$ we can write $A = \cup_{i=1}^N R_i$ for disjoint rectangles $R_i := (a_1^i, b_1^i] \times \cdots \times (a_n^i, b_n^i]$. We define for such A:

$$\mu(A) := \sum_{i=1}^n (b_1^i - a_1^i)(b_2^i - a_2^i) \cdots (b_n^i - a_n^i).$$

Note that the decomposition of A into rectangles is not unique, however one can verify that this is well defined and additive. If we can show that μ is countable additive, then we can apply Carathéodory's theorem to establish the existence of the Lebesgue measure.

Suppose that $(A_n)_{n=1}^\infty$ is a sequence of disjoint sets with $A_n \in \mathcal{A}_R$, such that $A = \cup_{i=1}^\infty A_i \in \mathcal{A}_R$. We wish to show that

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

Set $B_n = \cup_{i=n}^\infty A_i$, note $\cap_{i=1}^\infty B_i = \emptyset$ as the sets A_i are disjoint. Since $\mathcal{A}_R$ is a ring, $B_n = A \setminus \cup_{i=1}^{n-1} A_i \in \mathcal{A}_R$. By finite additivity of μ we have:

$$\mu(A) = \sum_{i=1}^{n-1} \mu(A_i) + \mu(B_n),$$

so it suffices to prove that $\mu(B_n) \to 0$ as $n \to \infty$. Suppose not, then there exists $\epsilon > 0$ such that $\mu(B_n) \geqslant 2\epsilon$ for all n. For each n we can find $C_n \in \mathcal{A}$ with $\overline{C}_n \subset B_n$ and $\mu(C_n \setminus B_n) \leqslant \epsilon 2^{-n}$. Then

$$\mu(B_n \setminus (C_1 \cap \cdots \cap C_n)) \leqslant \mu((B_1 \setminus C_1) \cup \ldots \cup (B_n \setminus C_n)) \leqslant \sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon.$$

Since $\mu(B_n) \geqslant 2\epsilon$, we must have $\mu(C_1 \cap \cdots \cap C_n) \geqslant \epsilon$, so $C_1 \cap \cdots \cap C_n \neq \emptyset$ and so $K_n = \overline{C}_1 \cap \cdots \cap \overline{C}_n \neq \emptyset$. Now, K_n is a nested sequence of non-empty compact sets, and so $\emptyset \neq \cap_{i=1}^{\infty} K_i \subset \cap_{i=1}^{\infty} B_i$ which is a contradiction.

Thus, we conclude that a Borel measure μ exists on $\mathbb{R}^n$ with the required property acting on rectangles. In order to establish uniqueness, we can invoke Corollary B.9, after noting that the set of rectangles is a π-system and that moreover we can write $\mathbb{R}^n$ as a countable disjoint union of rectangles, for example by taking the rectangles of the form $z + (0, 1]^n$, where $z \in \mathbb{Z}^n$. □

We note that the Lebesgue measure is translation invariant: $\mu(B + x) = \mu(B)$ for any $x \in \mathbb{R}^n$, $B \in \mathcal{B}(\mathbb{R}^n)$. To see this, for fixed $x \in \mathbb{R}^n$ let $\mu_x(B) = \mu(B)$. If B is a rectangle, then $\mu_x(R) = \mu(R)$ (since $b_1 - a_1 = (b_1 - x_1) - (a_1 - x_1)$, etc.) so by uniqueness $\mu_x = \mu$. We also note that Carathéodory's theorem actually shows us that the Lebesgue measure is actually defined on $\mathcal{M}$, a larger σ-algebra than $\mathcal{B}(\mathbb{R}^n)$. We call $\mathcal{M}$ the algebra of Lebesgue measurable sets. By construction, we have that the Lebesgue measure is complete when $\mathbb{R}^n$ is equipped with $\mathcal{M}$ as σ-algebra, however it is not complete on the Borel algebra. For any Lebesgue measurable subset $E \subset \mathbb{R}^n$ we can define the natural restriction of Lebesgue measure to E, which we also refer to as the Lebesgue measure.

Lemma B.12 (Borel regularity of Lebesgue measure)**.** *Suppose $A \in \mathcal{M}$ is Lebesgue measurable. Then for any $\epsilon > 0$ there exists an open set O and a closed set C such that $C \subset A \subset O$ and:*

$$\mu(O \setminus A) < \epsilon, \qquad \mu(A \setminus C) < \epsilon.$$

If $\mu(A) < \infty$, then we may take C to be compact.

Proof. First, let us assume $\mu(A) < \infty$. From the definition of Lebesgue measurability, we know that

$$\mu(A) = \mu^*(A) = \inf \sum_{n=1}^{\infty} \mu(A_n),$$

where the infimum is taken over all sequences $(A_n)_{n=1}^{\infty}$ of sets such that $A_n \in \mathcal{A}_R$ and $A \subset \cup_{i=1}^{n} A_n$. Since each $A_n \in \mathcal{A}_R$ is a finite disjoint union of rectangles, we may assume without loss of generality that each A_n is a rectangle. Fix $\epsilon > 0$. We can choose A_n such that:

$$\inf \sum_{n=1}^{\infty} \mu(A_n) < \mu(A) + \frac{\epsilon}{2}$$

For each rectangle A_n, we can find a rectangle $\tilde{A}_n$ with $A_n \subset \tilde{A}_n^\circ$ and $\mu(\tilde{A}_n) < \mu(A_n) + \frac{\epsilon}{2^{n+1}}$. Then let $O = \cup_{n=1}^\infty \tilde{A}_n^\circ$. By construction, $A \subset O$ and O is open. Moreover,

$$\mu(O) \leqslant \sum_{n=1}^\infty \mu(\tilde{A}_n) \leqslant \sum_{n=1}^\infty \mu(A_n) + \frac{\epsilon}{2}\sum_{n=1}^\infty 2^{-n} < \mu(A) + \epsilon.$$

We deduce that

$$\mu(O \setminus A) < \epsilon.$$

Now suppose $\mu(A) = \infty$. Set $A_k = A \cap \{|x| \leqslant k\}$, then $\mu(A_k) < \infty$, so we can find an open O_k with $\mu(O_k \setminus A_k) < \epsilon 2^{-k}$. We set $O = \cup_{k=1}^\infty O_k$. Then O is open and $A \subset O$. Moreover,

$$O \setminus A = (\cup_{k=1}^\infty O_k) \setminus A = \bigcup_{k=1}^\infty (O_k \setminus A) \subset \bigcup_{k=1}^\infty (O_k \setminus A_k)$$

so that

$$\mu(O \setminus A) \leqslant \sum_{k=1}^\infty \mu(O_k \setminus A_k) < \epsilon.$$

We have thus established the first part of the proof. For the second part, we note that if A is measurable, then so is A^c, and hence there exists an open O with $A^c \subset O$ and $\mu(O \setminus A^c) < \epsilon$. Set $C = O^c$. This is closed and $C \subset A$. Moreover, $A \setminus C = C^c \setminus A^c = O \setminus A^c$ so

$$\mu(A \setminus C) < \epsilon.$$

For the final observation, note that if $\mu(A) < \infty$, then since A_k is an increasing sequence with $\cup_k A_k = A$, we have that $\lim_{k\to\infty} \mu(A_k) = \mu(A) < \infty$, so there exists k such that $\mu(A \setminus A_k) = \mu(A) - \mu(A_k) < \frac{\epsilon}{2}$. Let $C \subset A_k$ be a closed set such that $\mu(A_k \setminus C) < \frac{\epsilon}{2}$. We have $\mu(A \setminus C) = \mu((A \setminus A_k) \cup (A_k \setminus C)) < \epsilon$, and moreover C is a subset of a bounded set, hence compact. ☐

We next show

Lemma B.13. *Let $A \subset \mathbb{R}^n$. Suppose that for any $\epsilon > 0$ there exists an open set O and a closed set C such that $C \subset A \subset O$ and:*

$$\mu(O \setminus C) < \epsilon.$$

Then $A = B_1 \cup N$, where $N \subset B_2$ where $B_1, B_2 \in \mathcal{B}(\mathbb{R}^n)$ with $\mu(B_2) = 0$.

Proof. For each i, we can find O_i open and C_i closed such that $C_i \subset A \subset O_i$ and

$$\mu(O_i \setminus C_i) < 2^{-i}.$$

We have that $B_1 = \cup_{i=1}^\infty C_i \in \mathcal{B}(\mathbb{R}^n)$ from the properties of σ-algebras. Furthermore, let $B_2 = \cap_{i=1}^\infty (O_i \setminus C_i)$. Again $B_2 \in \mathcal{B}(\mathbb{R}^n)$, and moreover:

$$\mu(B_2) \leqslant \mu\left(\cap_{i=1}^n (O_i \setminus C_i)\right) \leqslant 2^{-n+1}$$

for any n, so $\mu(B_2) = 0$. Since $A \setminus B_1 \subset B_2$ we are done. ☐

Now, noting that the union of a Borel set and a null set is Lebesgue measurable by the completeness of Lebesgue measure, we have established:

Theorem B.14. *Suppose $A \subset \mathbb{R}^n$. The following are equivalent:*

i) A is Lebesgue measurable.

ii) For any $\epsilon > 0$ there exists an open set O and a closed set C such that $C \subset A \subset O$ and:

$$\mu(O \setminus C) < \epsilon.$$

iii) $A = B_1 \cup N$, where $N \subset B_2$ where $B_1, B_2 \in \mathcal{B}(\mathbb{R}^n)$ with $\mu(B_2) = 0$.

B.2 Measurable functions

We next wish to introduce the idea of a *measurable function* between two measurable spaces. Suppose $(E, \mathcal{E})$ and $(G, \mathcal{G})$ are measurable spaces. We say $f : E \to G$ is *measurable* if $f^{-1}(A) \in \mathcal{E}$ whenever $A \in \mathcal{G}$. Note the similarity to the definition of continuous maps between topological spaces. If $(G, \mathcal{G}) = (\mathbb{R}, \mathcal{B})$, then we simply refer to a measurable function on $(E, \mathcal{E})$. If[3] $(G, \mathcal{G}) = (\mathbb{R}, \mathcal{B}[0, \infty])$, we refer to a non-negative measurable function. While convenient, this nomenclature has the slightly unfortunate consequence that a non-negative measurable function need not be a measurable function. If E is a topological space and $\mathcal{E} = \mathcal{B}(E)$, then a measurable function on $(E, \mathcal{E})$ is called a Borel function on E.

Exercise B.4. a) Suppose $(G, \mathcal{G})$ is a measurable space and E is any set. Show that if $f : E \to G$ is any function, the collection:

$$f^{-1}(\mathcal{G}) = \{f^{-1}(A) : A \in \mathcal{G}\},$$

is a σ-algebra, known as the pull-back σ-algebra.

b) Suppose $(E, \mathcal{E})$ and $(G, \mathcal{G})$ are measurable spaces, with $\mathcal{G} = \sigma(\mathcal{A})$ for some collection $\mathcal{A}$. Further suppose that $f : E \to G$ has the property that $f^{-1}(A) \in \mathcal{E}$ for all $A \in \mathcal{A}$. Show that

$$\{A \subset G : f^{-1}(A) \in \mathcal{E}\}$$

is a σ-algebra containing $\mathcal{A}$ and deduce that f is measurable.

c) Suppose $(E, \mathcal{E})$ is a measurable space. Show that $f : E \to \mathbb{R}$ is measurable if and only if

$$f^{-1}((-\infty, \lambda)) := \{x \in E : f(x) < \lambda\} \in \mathcal{E}, \quad \text{for all } \lambda \in \mathbb{R}.$$

and $f : [0, \infty]$ is measurable if and only if

$$f^{-1}([0, \lambda)) := \{x \in E : 0 \leqslant f(x) < \lambda\} \in \mathcal{E}, \quad \text{for all } 0 \leqslant \lambda < \infty.$$

[3] We give $[0, \infty]$ a topology by saying $U \subset [0, \infty]$ is open if and only if $\tan^{-1}(U)$ is open in the standard topology of $[0, \frac{\pi}{2}]$, where by convention $\tan(\pi/2) = +\infty$.

Exercise B.5. Suppose E, G are topological spaces equipped with their Borel σ-algebras.

a) Show that any continuous function $f : E \to G$ is measurable. Deduce that in particular any continuous function $f : E \to \mathbb{R}$ is a Borel function.

b) Show that if $g : G \to \mathbb{R}$ is continuous, and $f : E \to G$ is measurable then $g \circ f$ is measurable.

c) Let $G = \mathbb{R}^n$, with its canonical basis $(e_i)_{i=1}^n$. Show that $f : E \to G$ is measurable if and only if each component function $f_i = (f, e_i) : E \to \mathbb{R}$ is measurable.

An important feature of the class of measurable functions (and indeed a strong motivation for the development of the theory) is that it behaves well under limiting operations.

Theorem B.15. *Suppose $(E, \mathcal{E})$ is a measurable space and $(f_n)_{n=1}^\infty$ is a sequence of non-negative measurable functions. Then the functions $f_1 + af_2$ for $a \geqslant 0$ and $f_1 f_2$ are measurable, as are*

$$\inf_n f_n, \quad \sup_n f_n, \quad \liminf_n f_n, \quad \limsup_n f_n.$$

In particular, if $f_n(x) \to f(x)$, then f is measurable.

The same results hold for (not necessarily non-negative) measurable functions, provided the limiting functions are real valued (i.e. don't take the values $\pm\infty$).

Proof. By Exercise B.4 we know that $f_1^{-1}([0,\lambda))f_2^{-1}([0,\lambda)) \in \mathcal{E}$ for any $0 \leqslant \lambda < \infty$. Now, for any $0 \leqslant \lambda < \infty$:

$$(f_1 + af_2)^{-1}([0,\lambda]) = \bigcup_{r \in \mathbb{Q}, r \geqslant 0} [\{f < \lambda - ar\} \cap \{g < r\}] \in \mathcal{E}.$$

so $f_1 + af_2$ is measurable. We also note that f_1^2 is measurable, since $(f_1^2)^{-1}([0,\lambda)) = f_1^{-1}([0,\lambda^{\frac{1}{2}})) \in \mathcal{E}$. Combining these two results, and noting

$$f_1 f_2 = \frac{1}{4}\left((f_1 + f_2)^2 - (f_1 - f_2)^2\right)$$

we deduce that $f_1 f_2$ is measurable. Next, we note that

$$\{\inf_n f_n < \lambda\} = \bigcup_n \{f_n < \lambda\}$$

so $\inf_n f_n$ is measurable. Similarly,

$$\{\sup_n f_n < \lambda\} = \bigcup_{r \in \mathbb{Q}, r < \lambda} \left(\bigcap_n \{f_n < r\}\right)$$

so $\sup_n f_n$ is measurable. Finally, we note that $\limsup_n f_n = \inf_k g_k$, where $g_k = \sup_{n \geqslant k} f_n$ and $\liminf_n f_n = \sup_k h_k$, where $h_k = \inf_{n \geqslant k} f_n$. The last conclusion follows since if $f_n(x)$ converges, then $\lim_n f_n(x) = \limsup_n f_n(x) = \liminf_n f_n(x)$.

The proofs in the real valued case follow, mutatis mutandis. □

We establish one further result concerning measurable functions, before moving on to discuss the interaction between measurable functions and measures.

> **Exercise B.6.** Suppose $(E, \mathcal{E})$ is a measurable space and f is a measurable function on E. Show that the functions $f^+, f^-, |f|$ defined by:
>
> $$f^+(x) := \max\{f(x), 0\}, \qquad f^-(x) := \max\{-f(x), 0\}, \qquad |f| := f^+ + f^-,$$
>
> are non-negative measurable functions.

Theorem B.16 (Monotone Class Theorem)**.** *Let $(E, \mathcal{E})$ be a measurable space and let $\mathcal{A}$ be a π-system generating $\mathcal{E}$. Suppose V is a vector space of bounded functions $f : E \to \mathbb{R}$ such that:*

i) $1 \in V$ *and* $\mathbb{1}_A \in V$ *for all* $A \in \mathcal{A}$*;*

ii) *if* $f_n \in V$ *for all* n *and* f *is a bounded function such that* $0 \leqslant f_n \leqslant f_{n+1}$ *and* $f_n \to f$ *pointwise, then* $f \in V$*.*

Then V contains every bounded measurable function.

Proof. Let $\mathcal{D} = \{A \in \mathcal{E} : \mathbb{1}_A \in V\}$. Then the assumptions on V ensure $\mathcal{D}$ is a d-system containing $\mathcal{A}$, so $\mathcal{D} = \mathcal{E}$. Since V is a vector space, it must contain all finite linear combinations of indicator functions of measurable sets. If f is a bounded non-negative measurable function, then $f_n = 2^{-n}\lfloor 2^n f\rfloor$ is such a function, and moreover f_n is an increasing sequence which tends to f pointwise, so $f \in V$. Since any bounded measurable function can be written as the difference of two bounded non-negative measurable functions we're done. □

We shall now see how measurable functions interact with measures. Firstly, we note that a measurable function can be use to induce a measure on its image, given a measure on its domain. Suppose $(E, \mathcal{E})$ and $(G, \mathcal{G})$ are measurable spaces, μ is a measure on $(E, \mathcal{E})$ and $f : E \to G$ is a measurable function. We can define a measure on $(G, \mathcal{G})$, $f_*\mu$, called the push-forward or image measure by:

$$f_*\mu(A) = \mu(f^{-1}(A)), \qquad \text{for all } A \in \mathcal{G}.$$

Next we consider convergence in the context of a measurable space $(E, \mathcal{E}, \mu)$. Given some property P conditioned on a point $x \in E$, we say that P holds *almost everywhere* in E if

$$\mu\left(\{x \in E : P(x) \text{ is false}\}\right) = 0.$$

For example, we can consider $\mathbb{R}$ equipped with the Lebesgue measure, and introduce the Dirichlet function $f(x) = 1$ for $x \in \mathbb{Q}$, $f(x) = 0$ otherwise. Then we can say '$f = 0$ almost everywhere'. In circumstances where the choice of measure is ambiguous, one sometimes writes μ-almost everywhere. We often abbreviate almost everywhere to *a.e.*

If $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions on $(E, \mathcal{E}, \mu)$, we say $f_n \to f$ almost everywhere if

$$\mu(\{x \in E : f_n(x) \not\to f(x)\}) = 0.$$

Another notion of convergence that we can consider is *convergence in measure.* We say that $f_n \to f$ in measure if

$$\mu\left(\{x \in E : |f_n(x) - f(x)| > \epsilon\}\right) \to 0, \qquad \text{for all } \epsilon > 0.$$

The connection between these two notions is captured by

Theorem B.17. *Suppose $(E, \mathcal{E}, \mu)$ is a measure space, and $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions on E. Then:*

i) Suppose $\mu(E) < \infty$, then if $f_n \to f$ almost everywhere, then $f_n \to f$ in measure.

ii) If $f_n \to f$ in measure, then there exists a subsequence $(f_{n_k})_{k=1}^{\infty}$ such that $f_{n_k} \to f$ almost everywhere.

Proof. i) By considering $f_n - f$, assume wlog $f_n \to 0$ a.e.. Fix $\epsilon > 0$, then for any n:

$$\mu\left(\bigcap_{m \geqslant n} \{|f_m| \leqslant \epsilon\}\right) \leqslant \mu(\{|f_n| \leqslant \epsilon\})$$

Now, set $A_n = \bigcap_{m \geqslant n} \{|f_m| \leqslant \epsilon\}$. We have $A_n \subset A_{n+1}$ and

$$x \in \bigcup_n A_n \quad \iff \quad \text{there exists } N \text{ such that } |f_n(x)| \leqslant \epsilon \text{ for all } n \geqslant N.$$

Thus as $n \to \infty$, we have:

$$\mu(A_n) \geqslant \mu\left(\{x : f_n(x) \to 0\}\right) = \mu(E).$$

ii) Again, wlog suppose $f_n \to 0$ in measure. Set $n_1 = 1$. For each $k > 1$ we can find $n_k > n_{k+1}$ such that

$$\mu(\{|f_{n_k}| > 1/k\}) \leqslant 2^{-k}.$$

Now, let

$$A_k = \bigcup_{m \geqslant k} \{x \in E : |f_{n_m}(x)| > 1/m\}.$$

we have that $x \in \bigcap_k A_k$ if and only if for any k there exists $m \geqslant k$ such that $|f_{n_m}(x)| > 1/m$. Thus $x \notin \bigcap_k A_k$ if and only if there exists k such that for any $m \geqslant k$ we have $|f_{n_m}(x)| \leqslant 1/m$ and we conclude $f_{n_k} \to 0$ for all $x \notin \bigcap_k A_k$. Now, $A_{k+1} \subset A_k$ so, for any m:

$$\mu\left(\bigcap_k A_k\right) \leqslant \mu\left(A_m\right) = \mu\left(\bigcup_{m \geqslant k} \{|f_{n_m}(x)| > 1/m\}\right)$$
$$\leqslant \sum_{m \geqslant k} \mu(\{|f_{n_m}(x)| > 1/m\}) \leqslant 2^{m-1}$$

we conclude that $\mu\left(\bigcap_k A_k\right) = 0$ and thus $f_{n_k} \to 0$ a.e..

□

Exercise B.7. Let $E = [0,1]$ be equipped with the Lebesgue measure. Construct a sequence of functions $f_n : [0,1] \to [0,1]$ such that $f_n \to f$ in measure, but $(f_n(x))_{n=1}^{\infty}$ does not converge for any $x \in [0,1]$.

A final result concerns the measurability of a function which equals a measurable function almost everywhere.

Lemma B.18. *Let* $(E, \mathcal{E}, \mu)$ *be a* complete *measure space, and let* f *be a measurable function on* E. *If* $g : E \to \mathbb{R}$ *is such that* $f = g$ *almost everywhere, then* g *is measurable.*

Proof. Under the assumptions, $N = \{f \neq g\}$ is null, hence measurable by the completeness hypothesis, and $\mu(N) = 0$. Fix $a \in \mathbb{R}$. By assumption $A = \{f < a\}$ is measurable, and if we can show that $B = \{g < a\}$ is measurable then we will be done. Now, $B \cap A^c \subset N$, so by completeness $A \cap B^c$ is measurable, hence

$$B = A \cup (B \cap A^c)$$

is measurable. □

B.3 Integration

We now wish to define a notion of integration for measurable functions on some measure space $(E, \mathcal{E}, \mu)$. We approach this by first considering the case of non-negative measurable functions. These can be approximated from below by simple functions, which are finite linear combinations of characteristic functions on which the integral can be easily defined.

We say f is *simple* if

$$f = \sum_{n=1}^{k} \alpha_n \mathbb{1}_{A_n}$$

where $\alpha_n \in \mathbb{R}$ and $A_n \in \mathcal{E}$. For a non-negative simple function it is natural to define the integral as:

$$\mu(f) := \sum_{n=1}^{k} \alpha_n \mu(A_n)$$

Here, by convention $0 \cdot \infty = 0$. Alternative notations which we will make use of are:

$$\mu(f) = \int_E f d\mu = \int_E f(x) d\mu(x)$$

We note that α_n, A_n are not uniquely determined by f, however $\mu(f)$ is independent of the particular representation we choose.

Exercise B.8. a) Show that if $0 \leqslant \alpha_n, \beta_n < \infty$, $A_n, B_n \in \mathcal{E}$ satisfy

$$\sum_{n=1}^{k} \alpha_n \mathbb{1}_{A_n} = \sum_{n=1}^{l} \beta_n \mathbb{1}_{B_n},$$

then

$$\sum_{n=1}^{k} \alpha_n \mu(A_n) = \sum_{n=1}^{l} \beta_n \mu(B_n).$$

b) If f, g are simple functions and $a, b \geqslant 0$, show that

i) $\mu(af + bg) = a\mu(f) + b\mu(g)$.

ii) If $f \leqslant g$ then $\mu(f) \leqslant \mu(g)$.

iii) $f = 0$ a.e. if and only if $\mu(f) = 0$.

For a non-negative measurable function, we define the integral to be:

$$\mu(f) = \int_E f d\mu = \sup\{\mu(g) : g \text{ simple with } 0 \leqslant g \leqslant f\}.$$

By the results of Exercise B.8 this is consistent with the previous definition when f is simple. Note that $\mu(f)$ is permitted to take the value ∞. We also note that if f, g are non-negative measurable functions with $f \leqslant g$, then

$$\int_E f d\mu \leqslant \int_E g d\mu.$$

It also follows immediately from the definition that for any $\epsilon > 0$ there exists a simple function f_ϵ such that

$$\int_E |f - f_\epsilon| \, d\mu = \int_E (f - f_\epsilon) d\mu < \epsilon$$

To define the integral for functions which may take both positive and negative values, we first recall that if f is measurable then $f^+, f^-, |f|$ are non-negative measurable functions. We say that f is *integrable* if $\mu(|f|) < \infty$, in which case we define:

$$\mu(f) = \mu(f^+) - \mu(f^-).$$

Note that $f \leqslant g$ if and only if $f^+ \leqslant g^+$ and $f^- \geqslant g^-$, so that $f \leqslant g$ implies $\mu(f) \leqslant \mu(g)$. In particular, we have that $|\mu(f)| \leqslant \mu(|f|)$. By our comment above, for any $\epsilon > 0$ we can find a simple function f_ϵ such that

$$\int_E |f - f_\epsilon| \, d\mu < \epsilon$$

since we can approximate both f^+ and f^- by appropriate simple functions.

If at most one of $\mu(f^+)$ or $\mu(f^-)$ is infinite, then we can still define $\mu(f)$ by the same formula, but if both $\mu(f^+)$ and $\mu(f^-)$ are infinite then we can't sensibly assign a value to $\mu(f)$. We can also consider the case where f takes values in $\mathbb{R}^n$. In this case we pick a basis $(e_i)_{i=1}^n$ for $\mathbb{R}^n$ and write $f = \sum_{i=1}^n f_i e_i$. We say f is integrable if each f_i is integrable and we define:

$$\int_E f d\mu = \sum_{i=1}^n \left(\int_E f_i d\mu \right) e_i$$

This naturally gives a definition for functions taking complex values by the isomorphism $\mathbb{C} \simeq \mathbb{R}^2$.

B.3.1 Convergence theorems

A fundamental result in the Lebesgue theory of integration is the monotone convergence theorem, sometimes called Beppo Levi's Lemma. Suppose $(E, \mathcal{E}, \mu)$ is a measure space and that $(f_n)_{n=1}^{\infty}$ is a sequence of non-negative measurable functions which is increasing, i.e. $f_n(x) \leqslant f_{n+1}(x)$ for all $x \in E$ and $n \geqslant 1$. Then for each $x \in E$ the limit $f(x) = \lim_{n\to\infty} f_n(x)$ exists in $[0, \infty]$. We know that f is measurable, and the monotone convergence theorem asserts that $\mu(f) = \lim_{n\to\infty} \mu(f_n)$.

Theorem B.19 (Monotone convergence theorem)**.** *Let $(f_n)_{n=1}^{\infty}$ be an increasing sequence of non-negative integrable functions on a measure space $(E, \mathcal{E}, \mu)$ converging to f. Then*

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

Proof. Let $M = \sup_n \mu(f_n)$. We wish to show that $M = \mu(f)$. Since f_n is an increasing sequence, we have $f_n \leqslant f$ so that $\mu(f_n) \leqslant \mu(f)$. As this holds for all n, we deduce:

$$M \leqslant \mu(f) = \sup\{\mu(g) : g \text{ simple, } g \leqslant f\}.$$

If we can show that for any simple function g with $0 \leqslant g \leqslant f$ we have $\mu(g) \leqslant M$ then we're done. Suppose

$$g = \sum_{i=1}^{m} a_k \mathbb{1}_{A_k}$$

is such a function, where we may assume $A_k \in \mathcal{E}$ are disjoint without loss of generality. We define

$$g_n(x) = \min\{g(x), 2^{-n}\lfloor 2^n f_n \rfloor\}.$$

Then $(g_n)_{n=1}^{\infty}$ is an increasing sequence of simple functions, satisfying $g_n \leqslant f_n \leqslant f$ and $g_n \to g$. Fix $0 < \epsilon < 1$. Define the sets $A_{k,n}$ by

$$A_{k,n} = \{x \in A_k : g_n(x) \geqslant (1-\epsilon)a_k\}$$

Then since g_n is an increasing sequence, we have $A_{k,n} \subset A_{k,n+1}$. So by countable additivity we have $\mu(A_{k,n}) \to \mu(A_k)$ as $n \to \infty$. By construction we have

$$\mathbb{1}_{A_k} g_n \geqslant (1-\epsilon)a_k \mathbb{1}_{A_{k,n}}$$

so

$$\mu(\mathbb{1}_{A_k} g_n) \geqslant (1-\epsilon)a_k \mu(A_{k,n})$$

Now, noting that $g_n = \sum_{k=1}^{m} \mathbb{1}_{A_k} g_n$, and using the linearity result of Exercise B.8 we see

$$\mu(g_n) \geqslant (1-\epsilon)\sum_{k=1}^{m} a_k \mu(A_{k,n}) \to (1-\epsilon)\sum_{k=1}^{m} a_k \mu(A_k) = (1-\epsilon)\mu(g).$$

Now, $\mu(g_n) \leqslant \mu(f_n) \leqslant M$, so we have $(1-\epsilon)\mu(g) \leqslant M$ for any $\epsilon > 0$, hence $\mu(g) \leqslant M$. □

A straightforward corollary of this result is the following:

Corollary B.20. *Suppose $(f_n)_{n=1}^{\infty}$ is a sequence of non-negative measurable functions on a measure space $(E, \mathcal{E}, \mu)$. Then*

$$\int_E \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \left(\int_E f_n d\mu \right).$$

Another useful corollary:

Corollary B.21. *Suppose $(f_n)_{n=1}^{\infty}$ is a* decreasing *sequence of bounded measurable functions on a measure space $(E, \mathcal{E}, \mu)$. Then*

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

Proof. We take $g_n = f_1 - f_n$, since the f_n are bounded this is well defined (i.e we don't have to assign a value to $\infty - \infty$). Then $(g_n)_{n=1}^{\infty}$ is an increasing sequence and we can apply the usual monotone convergence theorem. □

> **Exercise($*$).** Give an example to show that Corollary B.21 fails if the boundedness assumption is dropped.

With the monotone convergence theorem in hand, we can readily show that the integral satisfies the properties we would expect.

Theorem B.22. *Suppose f, g are non-negative measurable functions on a measure space $(E, \mathcal{E}, \mu)$ and $a, b \geqslant 0$ are constants. Then:*

i) $\displaystyle\int_E (af + bg) d\mu = a \int_E f d\mu + b \int_E g d\mu$

ii) If $f \leqslant g$ then $\displaystyle\int_E f d\mu \leqslant \int_E g d\mu$

iii) $\displaystyle\int_E f d\mu = 0$ *if and only if $f = 0$ almost everywhere.*

Proof. Let

$$f_n(x) = \min\{2^n \lfloor 2^{-n} f(x) \rfloor, n\}, \qquad g_n(x) = \min\{2^n \lfloor 2^{-n} g(x) \rfloor, n\}.$$

Then $(f_n)_{n=1}^{\infty}, (g_n)_{n=1}^{\infty}$ is an increasing sequence of non-negative simple functions tending to f, g respectively, and clearly $(af_n + bg_n)_{n=1}^{\infty}$ is an increasing sequence tending to $af + bg$. Since these are simple functions, we have:

$$\int_E (af_n + bg_n) d\mu = a \int_E f_n d\mu + b \int_E g_n d\mu$$

and by the monotone convergence theorem, we can take the limit $n \to \infty$ to establish *i*). Point *ii*) we already noted follows directly from the definition of the integral.

Finally, if $f = 0$ almost everywhere, then we have $\mu(g) = 0$ for any simple $g \leqslant f$, and thus $\mu(f) = 0$. Now suppose $f(x) \neq 0$ almost everywhere. Then there exists $\epsilon > 0$ such that if $A = \{f \geqslant \epsilon\}$ then $\mu(A) > 0$. Then $g = \mathbb{1}_A \epsilon$ is a non-negative simple function with $g \leqslant f$ and $\mu(g) = \epsilon\mu(A) > 0$, hence $\mu(f) > 0$. □

We can extend this result to functions taking values in $\mathbb{R}$ as follows:

Theorem B.23. *Suppose f, g are integrable functions on a measure space $(E, \mathcal{E}, \mu)$ and $a, b \in \mathbb{R}$ are constants. Then:*

i) $\displaystyle\int_E (af + bg) d\mu = a \int_E f d\mu + b \int_E g d\mu$

ii) If $f \leqslant g$ then $\displaystyle\int_E f d\mu \leqslant \int_E g d\mu$

iii) If $f = 0$ almost everywhere then $\displaystyle\int_E f d\mu = 0.$

Proof. Note that it follows immediately from the definition that $\mu(-f) = -\mu(f)$. Suppose then that $a > 0$. We have:

$$\mu(af) = \mu(af^+) - \mu(af^-) = a\mu(f^+) - a\mu(f^-) = a\mu(f).$$

We also note that $(f+g)^+ - (f+g)^- = f + g = f^+ + g^+ - f^- - g^-$. As a consequence $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$, where both sides are sums of non-negative measurable functions, hence:

$$\mu((f+g)^+) + \mu(f^-) + \mu(g^-) = \mu((f+g)^-) + \mu(f^+) + \mu(g^+)$$

and on rearranging:

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

Combining our observations gives i). Noting that $f \leqslant g$ implies $0 \leqslant g - f$, we deduce $0 \leqslant \mu(g) - \mu(f)$ and thus ii) holds. Finally, if $f = 0$ almost everywhere, then $f^+, f^- = 0$ almost everywhere thus $\mu(f) = 0$. □

Suppose $(E, \mathcal{E}, \mu)$ is a measure space. If $A \in \mathcal{E}$ and f is integrable, then so is $f\mathbb{1}_A$. Recall also that A inherits a measure space structure in a natural way $(A, \mathcal{E}|_A, \mu|_A)$. It is relatively straightforward to see that $f|_A$ is integrable, and that we can define unambiguously

$$\int_A f d\mu := \int_E f\mathbb{1}_A d\mu = \int_A f|_A d\mu|_A,$$

By our linearity result i) above, if A, B are disjoint measurable sets, then

$$\int_A f d\mu + \int_B f d\mu = \int_{A\cup B} f d\mu.$$

We also note that by ii) we have that if $|f| \leqslant K$ almost everywhere and $\mu(E) < \infty$, then f is integrable and

$$\left| \int_E f d\mu \right| \leqslant K\mu(E).$$

A useful consequence of the monotone convergence theorem connects the Lebesgue integral to the Riemann integral.

Theorem B.24. *Let $A = (a_1, b_1] \times \cdots \times (a_n, b_n]$ be a rectangle in $\mathbb{R}^n$, and suppose $f : A \to \mathbb{R}$ is bounded. Then f is Riemann integrable*[4] *if and only if f is continuous almost everywhere. If so, f is integrable with respect to Lebesgue measure, μ, on A and moreover*

$$\mathscr{R}\int_A f(x)dx = \int_A f d\mu,$$

where $\mathscr{R}\int$ denotes the Riemann integral.

Proof. Since f is bounded, we may assume that $0 \leqslant f \leqslant K$ for some K, without loss of generality. We consider a sequence of partitions $\mathcal{P}_n$ of A such that $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$ and the mesh of $\mathcal{P}_n \to 0$. Correspondingly, we construct two sequences of functions, $\underline{f}_n, \overline{f}_n$ by

$$\underline{f}_n = \sum_{\pi \in \mathcal{P}_n} \inf_\pi f \mathbb{1}_\pi, \quad \overline{f}_n = \sum_{\pi \in \mathcal{P}_n} \sup_\pi f \mathbb{1}_\pi$$

which satisfy

$$0 \leqslant \underline{f}_n \leqslant \underline{f}_{n+1} \leqslant f \leqslant \overline{f}_{n+1} \leqslant \overline{f}_n \leqslant K.$$

Since each $\pi \in \mathcal{P}_n$ is a rectangle, it is certainly Lebesgue measurable and so $\underline{f}_n, \overline{f}_n$ are in fact simple functions. Moreover,

$$\int_A \underline{f}_n d\mu = L(f, \mathcal{P}_n), \qquad \int_A \overline{f}_n d\mu = U(f, \mathcal{P}_n),$$

where U, L are the usual upper and lower sums associated to a partition. The function f is Riemann integrable if and only if $L(f, \mathcal{P}_n), U(f, \mathcal{P}_n)$ have a common limit as $n \to \infty$, i.e:

$$U(f, \mathcal{P}_n) \to \mathscr{R}\int_A f(x)dx, \qquad L(f, \mathcal{P}_n) \to \mathscr{R}\int_A f(x)dx, \qquad \text{as } n \to \infty.$$

$(\underline{f}_n)_{n=1}^\infty$ is a monotone increasing sequence, bounded above by f, so there exists a bounded measurable function $\underline{f} \leqslant f$ such that $\underline{f} = \lim_{n\to\infty} \underline{f}_n = \sup \underline{f}_n$. Similarly, there exists a bounded measurable function $\overline{f} \geqslant f \geqslant \underline{f}$ such that $\overline{f} = \lim_{n\to\infty} \overline{f}_n = \inf \underline{f}_n$. By applying monotone convergence to $(\underline{f}_n)_{n=1}^\infty$ and $(\overline{f}_n)_{n=1}^\infty$ we have:

$$\lim_{n\to\infty} \int_A \underline{f}_n d\mu = \int_A \underline{f} d\mu \leqslant \int_A \overline{f} d\mu \leqslant \lim_{n\to\infty} \int_A \overline{f}_n d\mu$$

We deduce that f is Riemann integrable if and only if

$$\int_A \underline{f} d\mu = \int_A \overline{f} d\mu = \mathscr{R}\int_A f(x)dx.$$

This occurs if and only if $\underline{f} = \overline{f}$ almost everywhere.

We define the set of boundary points of $\mathcal{P}_n$ to be:

$$B_n = \bigcup_{\pi \in \mathcal{P}_n} \partial\pi \cap A$$

---

[4] For a discussion of the Riemann integral in $\mathbb{R}^n$ see Spivak: "Calculus on maniolds".

Clearly $\mu(B_n) = 0$, so the set $B = \cup_n B_n$ also has measure zero. Suppose $x \notin B$, then $\underline{f}(x) = \overline{f}(x)$ if and only if f is continuous at x. We conclude that f is Riemann integrable if and only if f is continuous almost everywhere. Since $\underline{f} \leqslant f \leqslant \overline{f}$ in this case, we deduce that f is almost everywhere equal to $\underline{f}$ and the result follows by Lemma B.18. □

This result, that Riemann integrability is equivalent to almost-everywhere continuity, is known as *Lebesgue's criterion for integrability.* In practice, many of the explicit integrals we encounter are Riemann integrals, and this gives us access to the standard toolkit to compute them. Where there's no possibility for ambiguity, we will often use the standard notation $\int dx$ or $\int d^n x$, etc. to denote Lebesgue integration.

The next convergence result for integrals we shall require allows us to drop the assumption that our sequence is monotone, but at the cost of a weakened result.

Lemma B.25 (Fatou's Lemma). *Suppose $(f_n)_{n=1}^{\infty}$ is a sequence of non-negative measurable functions on a measure space $(E, \mathcal{E}, \mu)$. Then*

$$\int_E \liminf_{n\to\infty} f_n d\mu \leqslant \liminf_{n\to\infty} \int_E f_n d\mu$$

Proof. Let $g_n = \inf_{m \geqslant n} f_m$. Then $(g_n)_{n=1}^{\infty}$ is an increasing sequence of non-negative measurable functions, which tends to $\liminf f_n$. Thus by monotone convergence

$$\int_E g_n d\mu \to \int_E \liminf_{n\to\infty} f_n d\mu.$$

On the other hand, for $k \geqslant n$ we have:

$$g_n \leqslant f_k,$$

hence

$$\int_E g_n d\mu \leqslant \int_E f_k d\mu \text{ for all } k \geqslant n \qquad \Longrightarrow \qquad \int_E g_n d\mu \leqslant \inf_{k\geqslant n} \int_E f_k d\mu.$$

Now, as $n \to \infty$

$$\inf_{k\geqslant n} \int_E f_k d\mu \to \liminf_{n\to\infty} \int_E f_n d\mu,$$

and we're done. □

> **Exercise(∗).** Construct a sequence $(f_n)_{n=1}^{\infty}$ of functions $f_n : [0,1] \to [0,\infty)$ satisfying the hypotheses of Fatou's Lemma such that the inequality is strict.

The next convergence result we shall establish is an especially useful one, and in particular will be invoked on many occasions during the course.

Theorem B.26 (The Dominated Convergence Theorem). *Suppose that $(E, \mathcal{E}, \mu)$ is a measure space and that $(f_n)_{n=1}^{\infty}$ is a sequence of measurable functions such that:*

i) There exists an integrable function g such that $|f_n| \leqslant g$.

ii) $f_n(x) \to f(x)$ for all x.

Then f is integrable and:

$$\int_E f_n d\mu \to \int_E f d\mu.$$

Proof. By Theorem B.15, f is measurable, and since $|f| \leqslant g$, we further have that f is integrable. We have

$$0 \leqslant g \pm f_n \to g \pm f,$$

so that $\liminf g \pm f_n = g \pm f$. By Fatou and properties of $\liminf, \limsup$ we have:

$$\int_E g d\mu + \int_E f d\mu = \int_E \liminf (g + f_n) d\mu \leqslant \liminf \int_E (g + f_n) d\mu = \int_E g d\mu + \liminf \int_E f_n d\mu$$
$$\int_E g d\mu - \int_E f d\mu = \int_E \liminf (g - f_n) d\mu \leqslant \liminf \int_E (g - f_n) d\mu = \int_E g d\mu - \limsup \int_E f_n d\mu$$

Rearranging, we have:

$$\int_E f d\mu \leqslant \liminf \int_E f_n d\mu \leqslant \limsup \int_E f_n d\mu \leqslant \int_E f d\mu,$$

hence

$$\liminf \int_E f_n d\mu = \limsup \int_E f_n d\mu = \int_E f d\mu,$$

and we're done. □

We note that the hypotheses can be weakened slightly: suppose the hypotheses hold almost everywhere, so that $X = \{x \in E : \text{ hypotheses fail}\}$ has measure zero, then by applying the Dominated Convergence Theorem to $f_n \mathbb{1}_{X^c}$, we can recover the same result.

Exercise B.9. Here μ is the Lebesgue measure on $\mathbb{R}$.

a) Show that $f : [0,1] \to \mathbb{R}$ given by $f(x) = \frac{1}{\sqrt{x}}$ is Lebesgue integrable, and that

$$\int_{[0,1]} f d\mu = \lim_{\epsilon \to 0} \mathscr{R} \int_\epsilon^1 \frac{1}{\sqrt{x}} dx.$$

b) Suppose $f : [0,1] \to \mathbb{R}$ is Riemann integrable on every interval $[\epsilon, 1]$, $\epsilon > 0$ and moreover

$$\mathscr{R} \int_\epsilon^1 |f(x)|\, dx \leqslant C$$

for some C independent of ϵ. Show that f is Lebesgue integrable with

$$\int_{[0,1]} f d\mu = \lim_{\epsilon \to 0} \mathscr{R} \int_\epsilon^1 f(x) dx.$$

c) Suppose $f : \mathbb{R} \to \mathbb{R}$ is Riemann integrable on every interval $[-R, R]$ and moreover

$$\mathscr{R} \int_{-R}^R |f(x)|\, dx \leqslant C$$

for some C independent of R. Show that f is Lebesgue integrable with

$$\int_{\mathbb{R}} f d\mu = \lim_{R\to\infty} \mathscr{R} \int_{-R}^{R} f(x)dx.$$

Give an example of a function such that

$$\lim_{R\to\infty} \mathscr{R} \int_{-R}^{R} f(x)dx$$

exists, but $f : \mathbb{R} \to \mathbb{R}$ is *not* Lebesgue integrable.

B.3.2 Product measures and Tonelli–Fubini

Given two measure spaces $(E, \mathcal{E}, \mu)$ and $(F, \mathcal{F}, \nu)$, we wish to construct a measure space on $E \times F$. We say a subset $E \times F$ is a rectangle if it is of the form $A \times B$, with $A \in \mathcal{E}$, $B \in \mathcal{F}$. We denote by $\mathcal{E} \boxtimes \mathcal{F}$ the collection of finite disjoint unions of rectangles. Note that if $A_i \in \mathcal{E}$, $B_i \in \mathcal{F}$ then

$$\begin{aligned}
&(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2), \\
&(A_1 \times B_1) \cup (A_2 \times B_2) = (A_1 \times B_1 \setminus B_2) \cup ((A_1 \cup A_2) \times (B_1 \cap B_2)) \cup (A_2 \times B_2 \setminus B_1)
\end{aligned}$$

where the right-hand side of the second line is a disjoint union of rectangles. Finally, since

$$(A_1 \times B_1)^c = (E \times B_1^c) \cup (A_1^c \times F),$$

we see that $\mathcal{E} \boxtimes \mathcal{F}$ is an algebra (hence a ring). We denote by $\mathcal{E} \otimes \mathcal{F}$ the σ-algebra generated by $\mathcal{E} \boxtimes \mathcal{F}$. We define a set function $\pi : \mathcal{E} \boxtimes \mathcal{F} \to [0, \infty]$ by

$$\pi\left(\bigcup_{i=1}^{N} (A_i \times B_i)\right) = \sum_{i=1}^{N} \mu(A_i)\nu(B_i),$$

where the rectangles $A_i \times B_i \in \mathcal{E} \times \mathcal{F}$, $i = 1, \ldots, N$, are assumed to be disjoint. Now suppose that $(A_j \times B_j)_{j=1}^{\infty}$ is a sequence of disjoint rectangles such that

$$\bigcup_{j=1}^{\infty} A_j \times B_j = A \times B \in \mathcal{E} \times \mathcal{F}.$$

We claim that

$$\mu(A)\mu(B) = \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j).$$

To see this, we note:

$$\mathbb{1}_A(x)\mathbb{1}_B(y) = \mathbb{1}_{A\times B}(x, y) = \sum_{j=1}^{\infty} \mathbb{1}_{A_j \times B_j}(x, y) = \sum_{j=1}^{\infty} \mathbb{1}_{A_j}(x)\mathbb{1}_{B_j}(y)$$

Integrating with respect to x, using Corollary B.20 we see

$$\mu(A)\mathbb{1}_B(y) = \sum_{j=1}^{\infty} \mathbb{1}_{B_j}(y) \int \mathbb{1}_{A_j} d\mu = \sum_{j=1}^{\infty} \mathbb{1}_{B_j}(y)\mu(A_j)$$

Integrating again with respect to y, by the same argument we find:

$$\mu(A)\mu(B) = \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j).$$

Note that this immediately implies that $\pi(C)$ is well defined for $C \in \mathcal{E} \boxtimes \mathcal{F}$, independent of how C is represented as a finite union of rectangles. We also note that $\mathcal{E} \boxtimes \mathcal{F}, \pi$ satisfy the conditions of Carathéodory's theorem, Theorem B.3, thus we can define an outer measure π^* on $E \times F$, whose restriction to $\mathcal{E} \otimes \mathcal{F}$ gives a measure, which agrees with π on $\mathcal{E} \boxtimes \mathcal{F}$. We call this measure on $\mathcal{E} \otimes \mathcal{F}$ the product measure, $\mu \times \nu$.

Note that the product measure $\mu \times \nu$ will not in general be unique. However, it will be if $(E, \mathcal{E}, \mu)$ and $(F, \mathcal{F}, \nu)$ are σ-finite. We say $(E, \mathcal{E}, \mu)$ is σ-finite if there exists a countable collection $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ of disjoint measurable sets, with $\mu(A_i) < \infty$, such that $E = \cup A_i$. If both $(E, \mathcal{E}, \mu)$ and $(F, \mathcal{F}, \nu)$ are σ-finite, then $\mathcal{E} \boxtimes \mathcal{F}$ satisfies the conditions to enable us to apply Corollary B.9 to deduce that $\mu \times \nu$ is the unique measure on $\mathcal{E} \otimes \mathcal{F}$ such that

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

A brief note of caution before we consider integration on product spaces. If E, F are topological spaces and $\mathcal{E}, \mathcal{F}$ are the Borel σ-algebras on their respective spaces, then $\mathcal{E} \otimes \mathcal{F}$ contains the Borel σ-algebra of $E \times F$ with the product topology. However, the two need not be equal in general. One important case where we do have equality is when E, F are σ-compact metric spaces[5]. In particular this is the case when $E = \mathbb{R}^n$, $F = \mathbb{R}^m$. By the uniqueness of Lebesgue measure, we have that the product measure restricted to $\mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^m) = \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^m)$ is the Lebesgue measure on $\mathbb{R}^{n+m}$.

We now wish to consider integration of a measurable function defined on $E \times F$. If $f : E \times F \to \mathbb{R}$, and $x \in E, y \in F$, we define the $x-$section, f_x and y-section, f^y as:

$$f_x(y) = f^y(x) = f(x, y).$$

We also introduce the x-section and y-section of a set $A \subset E \times F$ as:

$$A_x = \{y \in F : (x, y) \in A\}, \qquad A^y = \{x \in F : (x, y) \in A\}.$$

Note that $A_x \subset F$, $A^y \subset E$ and we have

$$(\mathbb{1}_A)_x = \mathbb{1}_{A_x}, \qquad (\mathbb{1}_A)^y = \mathbb{1}_{A^y},$$

Lemma B.27. *If $A \in \mathcal{E} \otimes \mathcal{F}$, then $A_x \in \mathcal{F}$ for all $x \in E$, $A^y \in \mathcal{E}$ for all $y \in F$. More generally, if f is $\mathcal{E} \otimes \mathcal{F}$-measurable, then f_x is $\mathcal{F}$-measurable and f^y is $\mathcal{E}$-measurable.*

[5] A topological space is σ-compact if it is the union of countably many compact sets.

Proof. Let

$$\mathcal{C} = \{A \subset E \times F : A_x \in \mathcal{F} \text{ for all } x \in E, A^y \in \mathcal{E} \text{ for all } y \in F\}$$

Certainly every measurable rectangle is in $\mathcal{C}$. We also have:

$$\left(\bigcup_{j=1}^{\infty} A_j\right)_x = \bigcup_{j=1}^{\infty} (A_j)_x\,, \qquad (A_x)^c = (A^c)_x,$$

and similarly for A^y, so that $\mathcal{C}$ is a σ-algebra and thus $\mathcal{E} \otimes \mathcal{F} \subset \mathcal{C}$. For the final part we note that for

$$(f_x)^{-1}(S) = \left(f^{-1}(S)\right)_x, \qquad (f^y)^{-1}(S) = \left(f^{-1}(S)\right)^y,$$

whence the result follows. □

Next we prove a special case of the Tonelli and Fubini theorems, where we restrict attention to the characteristic functions of a measurable set/

Lemma B.28. *Suppose μ, ν are finite, and let $A \in \mathcal{E} \otimes \mathcal{F}$. Then*

$$x \mapsto \nu(A_x), \qquad y \mapsto \mu(A^y)$$

are measurable functions, and

$$(\mu \times \nu)(A) = \int_E \nu(A_x) d\mu(x) = \int_E \mu(A^y) d\nu(y).$$

Proof. Let $\mathcal{C}$ consist of all sets $A \in \mathcal{E} \otimes \mathcal{F}$ for which the conclusion of the Lemma holds. Clearly $\mathcal{C}$ contains all rectangles, and these form a π-system. If we can show that $\mathcal{C}$ is a d-system, then we will be done by Lemma B.2.

Clearly $E \times F \in \mathcal{C}$. Suppose $A, B \in \mathcal{C}$ with $B \subset A$. Then$(A \setminus B)_x = A_x \setminus B_x$, so[6] $\nu((A \setminus B)_x) = \nu(A_x) - \nu(B_x)$, hence $x \mapsto \nu((A \setminus B)_x)$ is measurable, and

$$\begin{aligned}(\mu \times \nu)(A \setminus B) = (\mu \times \nu)(A) - (\mu \times \nu)(B) &= \int_E \nu(A_x) d\mu(x) - \int_E \nu(B_x) d\mu(x)\\ &= \int_E \nu((A \setminus B)_x) d\mu(x)\end{aligned}$$

A similar argument for $(A \setminus B)^y$ shows $A \setminus B \in \mathcal{C}$.

Now suppose $A_n \in \mathcal{C}$ with $A_n \subset A_{n+1}$ and let $A = \cup_n A_n$. Then by countable additivity we have $(x \mapsto \nu((A_n)_x))_{n=1}^{\infty}$ is a monotone increasing sequence of functions with limit $\nu(A_x)$. By monotone convergence we have $\nu(A_x)$ is measurable, with

$$\int_E \nu(A_x) d\mu(x) = \lim_{n\to\infty} \int_E \nu((A_n)_x) d\mu(x) = \lim_{n\to\infty} (\mu \times \nu)(A_n) = \mu \times \nu(A),$$

where in the final inequality we use countable additivity for $\mu \times \nu$. A similar argument for $\mu(A^y)$ establishes that $A \in \mathcal{C}$ and we're done. The extension to the case where μ, ν are assumed σ-finite is straightforward, and left as an exercise. □

[6] This is where the assumption that ν is finite is required

Exercise($*$). Show that Lemma B.28 holds if μ, ν are only assumed to be σ-finite.

We now prove two very closely related results, Tonelli's Theorem and Fubini's Theorem. They are often referred to together as the Tonelli-Fubini theorem.

Theorem B.29 (Tonelli)**.** *Assume $(E, \mathcal{E}, \mu)$, $(F, \mathcal{F}, \nu)$ are σ-finite measure spaces. Suppose $f : E \times F \to [0, \infty]$ is a non-negative measurable function. Then so are f_x, f^y, and setting:*

$$h(x) = \int_F f_x(y) d\nu(y), \qquad g(y) = \int_E f^y(x) d\mu(x),$$

we have that $h : E \to [0, \infty]$, $g : F \to [0, \infty]$ are non-negative measurable functions on their respective domains with:

$$\int_{E\times F} f d(\mu \times \nu) = \int_E h d\mu = \int_F g d\nu. \tag{B.2}$$

Proof. Take $(f_n)_{n=1}^{\infty}$ a monotone increasing sequence of non-negative simple functions with $f_n \to f$. Letting

$$h_n(x) = \int_F (f_n)_x(y) d\nu(y), \qquad g_n(y) = \int_E (f_n)^y(x) d\mu(x),$$

we have

$$\int_{E\times F} f_n d(\mu \times \nu) = \int_E h_n d\mu = \int_F g_n d\nu, \tag{B.3}$$

by the previous Lemma and the linearity of the integral. For each $x \in E$, we have that $((f_n)_x)_{n=1}^{\infty}$ is a monotone increasing sequence with $(f_n)_x \to f_x$ and similarly for $((f_n)^y)_{n=1}^{\infty}$. Thus, we have $(h_n)_{n=1}^{\infty}$ is an increasing sequence of functions with $h_n \to h$, and similarly for g_n by the monotone convergence theorem. Thus we can pass to the limit in (B.3) by the monotone convergence theorem to obtain (B.2). □

Theorem B.30 (Fubini)**.** *Assume $(E, \mathcal{E}, \mu)$, $(F, \mathcal{F}, \nu)$ are σ-finite measure spaces. Suppose $f : E \times F \to \mathbb{R}$ is an integrable function. Then $f_x : F \to \mathbb{R}$ is integrable for μ-almost every $x \in E$, as is $f^y : E \to \mathbb{R}$ for ν-almost every y. Thus*

$$h(x) = \int_F f_x(y) d\nu(y), \qquad g(y) = \int_E f^y(x) d\mu(x), \tag{B.4}$$

are defined almost everywhere. We have that $h : E \to \mathbb{R}$, $g : F \to \mathbb{R}$ are integrable functions on their respective domains, and:

$$\int_{E\times F} f d(\mu \times \nu) = \int_E h d\mu = \int_F g d\nu. \tag{B.5}$$

Proof. Write $f = f^+ - f^-$, with $f^\pm$ non-negative and integrable. By Tonelli applied to $f^\pm$ we find $h^\pm$, $g^\pm$ such that

$$\int_{E\times F} f^\pm d(\mu \times \nu) = \int_E h^\pm d\mu = \int_F g^\pm d\nu.$$

The first integral is finite by assumption, so we must have that $h^\pm, g^\pm$ are finite almost everywhere, and moreover are integrable. Thus $h = h^+ - h^-$ and $g = g^+ - g^-$ satisfy (B.4) and we also deduce (B.5). $\square$

In the particular case where $E = \mathbb{R}^n$, $F = \mathbb{R}^m$ equipped with their Borel sets and Lebesgue measure, then we conclude that if $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with

$$\int_{\mathbb{R}^{n+m}} |f(x,y)|\, dxdy < \infty$$

then

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x,y)dy \right) dx = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x,y)dx \right) dy = \int_{\mathbb{R}^{n+m}} f(z)dz$$

with obvious notation.

In combination, Tonelli–Fubini together with the Dominated Convergence Theorem are very powerful, and typically suffice for the majority of convergence related results that we require in standard analysis.

Exercise B.10. Let $(f_n)_{n=1}^\infty$ be a sequence of measurable functions $f_n : \mathbb{R}^m \to \mathbb{R}$ such that

$$\sum_{n=1}^\infty \int_{\mathbb{R}^m} |f_n|\, dx < \infty.$$

Show that:

$$f(x) = \sum_{n=1}^\infty f_n(x)$$

converges for a.e. $x \in \mathbb{R}^m$, and

$$\int_{\mathbb{R}^m} f dx = \sum_{n=1}^\infty \int_{\mathbb{R}^m} f_n dx.$$

B.4 The L^p-spaces

Given a measure space $(E, \mathcal{E}, \mu)$, we say that a measurable complex-valued[7] function f belongs to $\mathcal{L}^p(E, \mu)$ for some $p \leqslant 1 < \infty$ if

$$\|f\|_{L^p} := \left(\int_E |f|^p \, d\mu \right)^{\frac{1}{p}} = (\mu(|f|^p))^{\frac{1}{p}} < \infty.$$

We say that $f \in \mathcal{L}^\infty(E, \mu)$ if f is bounded almost everywhere, that is there exists $0 \leqslant K < \infty$ such that

$$\mu(\{|f(x)| > K\}) = 0.$$

If so, then we define

$$\|f\|_{L^\infty} := \inf\{K : \mu(\{|f(x)| > K\}) = 0\}.$$

We can show:

[7] We can also assume f is real-valued

Lemma B.31. *For* $1 \leqslant p \leqslant \infty$*, the function* $f \mapsto \|f\|_{L^p}$ *defines a* seminorm *on* $\mathcal{L}^\infty(E,\mu)$*. That is:*

i) $\|\cdot\|_{L^p}$ *is non-negative:*

$$\|f\|_{L^p} \geqslant 0, \qquad \textit{for all } f \in \mathcal{L}^p(E,\mu).$$

ii) $\|\cdot\|_{L^p}$ *is homogeneous:*

$$\|\lambda f\|_{L^p} = |\lambda| \, \|f\|_{L^p}, \qquad \textit{for all } f \in \mathcal{L}^p(E,\mu), \lambda \in \mathbb{C}.$$

iii) $\|\cdot\|_{L^p}$ *satisfies the triangle inequality:*

$$\|f+g\|_{L^p} \leqslant \|f\|_{L^p} + \|g\|_{L^p}, \qquad \textit{for all } f, g \in \mathcal{L}^p(E,\mu).$$

Proof. See Exercise B.11. □

We note that the crucial property that is missing and prevents $\mathcal{L}^p(E,\mu)$ from being a normed space is *positivity*, i.e. that $\|f\|_{L^p} = 0$ if and only if $f = 0$. By Theorem B.22, we know that $\|f\|_{L^p} = 0$ if and only if $f = 0$ holds *almost everywhere*. In order to construct a normed space, we must quotient out the elements of $\mathcal{L}^p(E,\mu)$ which satisfy $\|f\|_{L^p} = 0$. To do this, we introduce an equivalence relation according to:

$$f \sim g \qquad \Longleftrightarrow \qquad f - g = 0 \text{ a.e.}$$

It is straightforward to see that $\sim$ defines an equivalence relation on $\mathcal{L}^p(E,\mu)$ and moreover, by the reverse triangle inequality

$$f \sim g \qquad \Longrightarrow \qquad \|f\|_{L^p} = \|g\|_{L^p}.$$

Thus we can define a new space

$$L^p(E,\mu) = \mathcal{L}^p(E,\mu)/\sim,$$

and $\|\cdot\|_{L^p}$ descends to a norm on the quotient space by:

$$\|[f]_\sim\|_{L^p} := \|f\|_{L^p}.$$

In practice, we usually elide the distinction between the function $f \in \mathcal{L}^p(E,\mu)$ and the equivalence class of functions $[f]_\sim \in L^p(E,\mu)$, so it is standard to speak of a function f belonging to $L^p(E,\mu)$. One should always remember, however, that in general statements about elements of $L^p(E,\mu)$ hold at most almost everywhere. It is immediate that we have

Lemma B.32. *The space* $L^p(E,\mu)$*, equipped with the norm* $\|\cdot\|_{L^p}$*, is a normed vector space.*

In the case where $E = \mathbb{R}^n$ equipped with the σ-algebra of Lebesgue measurable sets and the Lebesgue measure, we typically write $L^p(\mathbb{R}^n)$ instead of $L^p(\mathbb{R}^n, dx)$ to denote the associated spaces.

Exercise B.11. Let $(E, \mathcal{E}, \mu)$ be a measure space. Show that $\|\cdot\|_{L^p}$ defines a seminorm on $\mathcal{L}^p(E, \mu)$ for $1 \leqslant p \leqslant \infty$:

a) First check that the homogeneity and non-negativity properties are satisfied.

b) Establish the triangle inequality for the special cases $p = 1, \infty$.

c) Next prove Young's inequality: if $a, b \in \mathbb{R}_+$ and $p, q > 1$ with $p^{-1} + q^{-1} = 1$ then:

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

Hint: set $t = p^{-1}$, *consider the function* $\log\left[ta^p + (1-t)b^q\right]$ *and use the concavity of the logarithm*

d) With[8] $p, q \geqslant 1$ such that $p^{-1} + q^{-1} = 1$, show that if $\|f\|_{L^p} = 1$ and $\|g\|_{L^q} = 1$ then

$$\int_E |fg| \, d\mu \leqslant 1$$

Deduce Hölder's inequality:

$$\int_E |fg| \, d\mu \;\leqslant \|f\|_{L^p} \|g\|_{L^q}\,, \qquad \text{for all } f \in \mathcal{L}^p(E, \mu), \;\; y \in \mathcal{L}^q(E, \mu).$$

e) Show that if $f, g \in \mathcal{L}^p(E, \mu)$

$$\|f + g\|_{L^p}^p \leqslant \int_E |f| \, |f + g|^{p-1} \, d\mu + \int_E |g| \, |f + g|^{p-1} \, d\mu$$

Apply Hölder's inequality to deduce:

$$\|f + g\|_{L^p}^p \leqslant \left(\|f\|_{L^p} + \|g\|_{L^p}\right) \|f + g\|_{L^p}^{p-1}$$

and conclude

$$\|f + g\|_{L^p} \leqslant \|f\|_{L^p} + \|g\|_{L^p}\,.$$

This is Minkowski's inequality.

Exercise B.12. a) Suppose that $\mu(E) < \infty$. Show that if $f \in \mathcal{L}^p(E, \mu)$, then $f \in \mathcal{L}^q(E, \mu)$ for any $1 \leqslant q < p$, with

$$\|f\|_{L^q} \leqslant \mu(E)^{\frac{p-q}{qp}} \|f\|_{L^p}\,.$$

b) Suppose that $f \in \mathcal{L}^{p_0}(E, \mu) \cap \mathcal{L}^{p_1}(E, \mu)$ with $p_0 < p_1 \leqslant \infty$. For $0 \leqslant \theta \leqslant 1$, define p_θ by

$$\frac{1}{p_\theta} = \frac{1 - \theta}{p_0} + \frac{\theta}{p_1}.$$

Show that $f \in \mathcal{L}^{p_\theta}(E, \mu)$ with

$$\|f\|_{L^{p_\theta}} \leqslant \|f\|_{L^{p_0}}^{1-\theta} \|f\|_{L^{p_1}}^{\theta}\,.$$

[8] We permit p, q to take the value ∞ with the convention $\infty^{-1} = 0$

B.4.1 Completeness

The most important property of the L^p-spaces is that in fact, they are complete, i.e. they are Banach spaces. To establish this, we first prove the following result.

Lemma B.33. *Suppose $(E, \mathcal{E}, \mu)$ is a measure space and $1 \leqslant p < \infty$. Let $(g_n)_{n=1}^{\infty}$ be a sequence with $g_n \in L^p(E, \mu)$ such that*

$$\sum_{n=1}^{\infty} \|g_n\|_{L^p} < \infty$$

then there exists $f \in L^p(E, \mu)$ such that

$$\sum_{n=1}^{\infty} g_n = f,$$

where the sum converges pointwise almost everywhere, and in $L^p(E, \mu)$.

Proof. Fix representative[9] functions $\tilde{g}_n \in \mathcal{L}^p(E, \mu)$ corresponding to each $g_n \in L^p(E, \mu)$. Define $h_n, h : E \to [0, \infty]$ by

$$h_n = \sum_{k=1}^{n} |\tilde{g}_k|, \qquad h = \sum_{k=1}^{\infty} |\tilde{g}_k|.$$

Note that $(h_n)_{n=1}^{\infty}$ is a monotone increasing sequence of non-negative measurable functions, converging pointwise to h, so by the monotone convergence theorem we have

$$\int_E h^p d\mu = \lim_{n\to\infty} \int_E h_n^p d\mu.$$

By Minkowski's inequality we see

$$\|h_n\|_{L^p} \leqslant \sum_{k=1}^{n} \|g_k\|_{L^p} \leqslant K =: \sum_{k=1}^{\infty} \|g_k\|_{L^p}.$$

It follows that $h \in \mathcal{L}^p(E, \mu)$ with $\|h\|_{L^p} \leqslant K$, which in particular implies that h is finite almost everywhere. At each point x such that $h(x) < \infty$, we have that $\sum_{k=1}^{\infty} \tilde{g}_k(x)$ converges absolutely, hence converges by the completeness of $\mathbb{C}$. We deduce that $\sum_{k=1}^{\infty} \tilde{g}_k$ converges pointwise almost everywhere and we define:

$$f(x) = \begin{cases} \sum_{k=1}^{\infty} \tilde{g}_k(x) & \text{if the sum converges} \\ 0 & \text{otherwise} \end{cases}$$

Now, we have that $|f| \leqslant h$, which implies $\|f\|_{L^p} \leqslant \|h\|_{L^p} \leqslant K$, and moreover:

$$\left| f - \sum_{k=1}^{n} \tilde{g}_k \right|^p \leqslant \left(|f| + \sum_{k=1}^{n} |\tilde{g}_k| \right)^p \leqslant (2h)^p.$$

[9] We typically don't state this point explicitly, but on this occasion we will make the distinction between $\mathcal{L}^p$ and L^p

Since h^p is integrable, by the Dominated Convergence Theorem (Thm B.26) we deduce that

$$\int_E \left| f - \sum_{k=1}^{n} \tilde{g}_k \right|^p d\mu \to 0, \qquad \text{as } n \to \infty,$$

which implies $\sum_{k=1}^{\infty} \tilde{g}_k$ converges to f in L^p. Noting that a different choice of representatives $\tilde{g}_n$ will result in h_n, h, f which differ from those defined above only on a set of measure zero, since the union of countably many sets of measure zero also has measure zero we are done. □

With this result in hand, we are able to establish

Theorem B.34 (Riesz-Fischer Theorem). *Suppose $(E, \mathcal{E}, \mu)$ is a measure space and $1 \leqslant p \leqslant \infty$. Then $L^p(E, \mu)$ is complete.*

Proof. To prove completeness, suppose $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence with respect to the L^p-norm. It suffices to show there exists $f \in L^p(E, \mu)$ with $f_n \to f$ in L^p. We split the cases $p < \infty$ and $p = \infty$.

1. First suppose $1 \leqslant p < \infty$. Then by the Cauchy property we can find a subsequence $(f_{n_k})_{k=1}^{\infty}$ such that

   $$\left\| f_{n_{k+1}} - f_{n_k} \right\|_{L^p} < 2^{-k}.$$

   Set $g_k = f_{n_{k+1}} - f_{n_k}$. By construction we have:

   $$\sum_{k=1}^{\infty} \|g_k\|_{L^p} < \sum_{k=1}^{\infty} 2^{-k} = 1,$$

   so by Lemma B.33 there exists $g \in L^p(E, \mu)$ such that

   $$\sum_{k=1}^{\infty} g_k = g$$

   with the sum converging pointwise a.e. and in L^p. Noting that $f_{n_{j+1}} = f_{n_1} + \sum_{k=1}^{j} g_k$, we deduce that $(f_{n_k})_{k=1}^{\infty}$ converges in L^p to some $f \in L^p(E, \mu)$. It follows by a standard argument using the fact it is a Cauchy sequence that $(f_n)_{n=1}^{\infty}$ converges to f in L^p.

2. Now consider the case $p = \infty$. Since (f_n) is Cauchy in $L^\infty(E, \mu)$, for each $m \in \mathbb{N}$ there exists n such that for any $j, m \geqslant n$ we have

   $$|f_j(x) - f_k(x)| < \frac{1}{m} \qquad \text{for all } x \in N_{j,k,m}^c$$

   where $\mu(N_{j,k,m}) = 0$. Let

   $$N = \bigcup_{j,k,m} N_{j,k,m},$$

then $\mu(N) = 0$ and further we have that for any $m \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that for $j, k \geqslant n$:

$$\sup_{x \in N^c} |f_j(x) - f_k(x)| < \frac{1}{m}, \tag{B.6}$$

so by the completeness of $\mathbb{C}$ we have that for each $x \in N^c$ we have $f_j(x) \to f(x)$ for some $f(x) \in \mathbb{C}$ We let $f(x) = 0$ for $x \in N$, so that $f : E \to \mathbb{C}$. Sending $k \to \infty$ in (B.6) we see that for $j \geqslant n$

$$\sup_{x \in N^c} |f_j(x) - f(x)| < \frac{1}{m},$$

whence we conclude that $\|f\|_{L^\infty} < \infty$ and $f_j \to f$ in L^∞. □

Note that we have in fact proved the stronger result:

Corollary B.35. *Suppose $(f_n)_{n=1}^\infty$ is a Cauchy sequence in $L^p(E, \mu)$ for $1 \leqslant p \leqslant \infty$. Then there exists a subsequence $(f_{n_k})_{k=1}^\infty$ which converges pointwise almost everywhere.*

B.4.2 Density

It is often useful, when discussing topological spaces to identify dense subsets consisting of 'nice' or 'concrete' objects, for example elements of $\mathbb{Q}$ can be easily discussed, while a general element of $\mathbb{R}$ is typically expressible only as some limit of elements of $\mathbb{Q}$. In the main body of the course we shall establish that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$. For a general measure space, we don't necessarily have a notion of continuity or smoothness, but we can show

Theorem B.36. *Let S be the set of all complex, measurable, simple functions on E such that:*

$$\mu(\{x : s(x) \neq 0\}) < \infty.$$

Then S is dense in $L^p(E, \mu)$ for $1 \leqslant p < \infty$.

Proof. Clearly $S \subset L^p(E, \mu)$. Now, suppose $f \geqslant 0$ with $f \in L^p(E, \mu)$ and let

$$f_n(x) = \min\{2^n \lfloor 2^{-n} f(x) \rfloor, n\}.$$

We have $f_n \in S$ and $0 \leqslant f_n \leqslant f$, so that $f_n \in L^p(E, \mu)$. Further, we know that $f_n(x) \to f(x)$ and moreover

$$|f - f_n|^p \leqslant |f|^p,$$

so by the Dominated convergence Theorem (Thm B.26) we deduce

$$\int_E |f - f_n|^p \, d\mu \to 0$$

hence $f_n \to f$ in L^p. A general (i.e. complex valued) element of $L^p(E, \mu)$ may be written as:

$$f = f_r^+ - f_r^- + i(f_i^+ - f_i^-),$$

where $f_r^\pm, f_i^\pm$ are non-negative elements of $L^p(E, \mu)$, hence the result follows. □