Part V

Lebesgue Integration Theory

17

Introduction: What are measures and why "measurable" sets

Definition 17.1 (Preliminary). *A* ***measure*** μ *"on" a set* X *is a function* $\mu : 2^X \to [0, \infty]$ *such that*

1. $\mu(\emptyset) = 0$
2. *If* $\{A_i\}_{i=1}^N$ *is a finite* ($N < \infty$) *or countable* ($N = \infty$) *collection of subsets of* X *which are pair-wise disjoint (i.e.* $A_i \cap A_j = \emptyset$ *if* $i \neq j$*) then*

$$\mu(\cup_{i=1}^N A_i) = \sum_{i=1}^N \mu(A_i).$$

Example 17.2. Suppose that X is any set and $x \in X$ is a point. For $A \subset X$, let

$$\delta_x(A) = \begin{cases} 1 \text{ if } & x \in A \\ 0 \text{ if } & x \notin A. \end{cases}$$

Then $\mu = \delta_x$ is a measure on X called the Dirac delta measure at x.

Example 17.3. Suppose that μ is a measure on X and $\lambda > 0$, then $\lambda \cdot \mu$ is also a measure on X. Moreover, if $\{\mu_\alpha\}_{\alpha \in J}$ are all measures on X, then $\mu = \sum_{\alpha \in J} \mu_\alpha$, i.e.

$$\mu(A) = \sum_{\alpha \in J} \mu_\alpha(A) \text{ for all } A \subset X$$

is a measure on X. (See Section 2 for the meaning of this sum.) To prove this we must show that μ is countably additive. Suppose that $\{A_i\}_{i=1}^\infty$ is a collection of pair-wise disjoint subsets of X, then

$$\begin{aligned} \mu(\cup_{i=1}^\infty A_i) &= \sum_{i=1}^\infty \mu(A_i) = \sum_{i=1}^\infty \sum_{\alpha \in J} \mu_\alpha(A_i) \\ &= \sum_{\alpha \in J} \sum_{i=1}^\infty \mu_\alpha(A_i) = \sum_{\alpha \in J} \mu_\alpha(\cup_{i=1}^\infty A_i) \\ &= \mu(\cup_{i=1}^\infty A_i) \end{aligned}$$

wherein the third equality we used Theorem 4.22 and in the fourth we used that fact that μ_α is a measure.

Example 17.4. Suppose that X is a set $\lambda : X \to [0,\infty]$ is a function. Then

$$\mu := \sum_{x\in X} \lambda(x)\delta_x$$

is a measure, explicitly

$$\mu(A) = \sum_{x\in A} \lambda(x)$$

for all $A \subset X$.

17.1 The problem with Lebesgue "measure"

So far all of the examples of measures given above are "counting" type measures, i.e. a weighted count of the number of points in a set. We certainly are going to want other types of measures too. In particular, it will be of great interest to have a measure on $\mathbb{R}$ (called Lebesgue measure) which measures the "length" of a subset of $\mathbb{R}$. Unfortunately as the next theorem shows, there is no such reasonable measure of length if we insist on measuring all subsets of $\mathbb{R}$.

Theorem 17.5. *There is no measure* $\mu : 2^{\mathbb{R}} \to [0,\infty]$ *such that*

1. $\mu([a,b)) = (b-a)$ *for all* $a < b$ *and*
2. *is translation invariant, i.e.* $\mu(A+x) = \mu(A)$ *for all* $x \in \mathbb{R}$ *and* $A \in 2^{\mathbb{R}}$, *where*
$$A + x := \{y + x : y \in A\} \subset \mathbb{R}.$$
In fact the theorem is still true even if (1) is replaced by the weaker condition that $0 < \mu((0,1]) < \infty$.

The counting measure $\mu(A) = \#(A)$ is translation invariant. However $\mu((0,1]) = \infty$ in this case and so μ does not satisfy condition 1.

Proof. First proof. Let us identify $[0,1)$ with the unit circle $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ by the map

$$\phi(t) = e^{i2\pi t} = (\cos 2\pi t + i \sin 2\pi t) \in S^1$$

for $t \in [0,1)$. Using this identification we may use μ to define a function ν on 2^{S^1} by $\nu(\phi(A)) = \mu(A)$ for all $A \subset [0,1)$. This new function is a measure on S^1 with the property that $0 < \nu((0,1]) < \infty$. For $z \in S^1$ and $N \subset S^1$ let

$$zN := \{zn \in S^1 : n \in N\}, \tag{17.1}$$

that is to say $e^{i\theta}N$ is N rotated counter clockwise by angle θ. We now claim that ν is invariant under these rotations, i.e.

$$\nu(zN) = \nu(N) \tag{17.2}$$

for all $z \in S^1$ and $N \subset S^1$. To verify this, write $N = \phi(A)$ and $z = \phi(t)$ for some $t \in [0,1)$ and $A \subset [0,1)$. Then

$$\phi(t)\phi(A) = \phi(t + A \operatorname{mod} 1)$$

where for $A \subset [0,1)$ and $\alpha \in [0,1)$,

$$\begin{aligned} t + A \operatorname{mod} 1 &:= \{a + t \operatorname{mod} 1 \in [0,1) : a \in N\} \\ &= (a + A \cap \{a < 1-t\}) \cup ((t-1) + A \cap \{a \geq 1-t\}). \end{aligned}$$

Thus

$$\begin{aligned} \nu(\phi(t)\phi(A)) &= \mu(t + A \operatorname{mod} 1) \\ &= \mu\left((a + A \cap \{a < 1-t\}) \cup ((t-1) + A \cap \{a \geq 1-t\})\right) \\ &= \mu\left((a + A \cap \{a < 1-t\})\right) + \mu\left(((t-1) + A \cap \{a \geq 1-t\})\right) \\ &= \mu\left(A \cap \{a < 1-t\}\right) + \mu\left(A \cap \{a \geq 1-t\}\right) \\ &= \mu\left((A \cap \{a < 1-t\}) \cup (A \cap \{a \geq 1-t\})\right) \\ &= \mu(A) = \nu(\phi(A)). \end{aligned}$$

Therefore it suffices to prove that no finite non-trivial measure ν on S^1 such that Eq. (17.2) holds. To do this we will "construct" a non-measurable set $N = \phi(A)$ for some $A \subset [0,1)$. Let

$$R := \{z = e^{i2\pi t} : t \in \mathbb{Q}\} = \{z = e^{i2\pi t} : t \in [0,1) \cap \mathbb{Q}\}$$

– a countable subgroup of S^1. As above R acts on S^1 by rotations and divides S^1 up into equivalence classes, where $z, w \in S^1$ are equivalent if $z = rw$ for some $r \in R$. Choose (using the axiom of choice) one representative point n from each of these equivalence classes and let $N \subset S^1$ be the set of these representative points. Then every point $z \in S^1$ may be uniquely written as $z = nr$ with $n \in N$ and $r \in R$. That is to say

$$S^1 = \coprod_{r \in R} (rN) \tag{17.3}$$

where $\coprod_\alpha A_\alpha$ is used to denote the union of pair-wise disjoint sets $\{A_\alpha\}$. By Eqs. (17.2) and (17.3),

$$\nu(S^1) = \sum_{r \in R} \nu(rN) = \sum_{r \in R} \nu(N).$$

The right member from this equation is either 0 or ∞, 0 if $\nu(N) = 0$ and ∞ if $\nu(N) > 0$. In either case it is not equal $\nu(S^1) \in (0,1)$. Thus we have reached the desired contradiction. ■

Proof. Second proof of Theorem 17.5. For $N \subset [0,1)$ and $\alpha \in [0,1)$, let

$$\begin{aligned} N^\alpha &= N + \alpha \bmod 1 \\ &= \{a + \alpha \bmod 1 \in [0,1) : a \in N\} \\ &= (\alpha + N \cap \{a < 1 - \alpha\}) \cup ((\alpha - 1) + N \cap \{a \geq 1 - \alpha\}). \end{aligned}$$

Then

$$\begin{aligned} \mu(N^\alpha) &= \mu(\alpha + N \cap \{a < 1 - \alpha\}) + \mu((\alpha - 1) + N \cap \{a \geq 1 - \alpha\}) \\ &= \mu(N \cap \{a < 1 - \alpha\}) + \mu(N \cap \{a \geq 1 - \alpha\}) \\ &= \mu(N \cap \{a < 1 - \alpha\} \cup (N \cap \{a \geq 1 - \alpha\})) \\ &= \mu(N). \end{aligned} \tag{17.4}$$

We will now construct a bad set N which coupled with Eq. (17.4) will lead to a contradiction. Set

$$Q_x := \{x + r \in \mathbb{R} : r \in \mathbb{Q}\} = x + \mathbb{Q}.$$

Notice that $Q_x \cap Q_y \neq \emptyset$ implies that $Q_x = Q_y$. Let $\mathcal{O} = \{Q_x : x \in \mathbb{R}\}$ – the orbit space of the $\mathbb{Q}$ action. For all $A \in \mathcal{O}$ choose $f(A) \in [0, 1/3) \cap A$[1] and define $N = f(\mathcal{O})$. Then observe:

1. $f(A) = f(B)$ implies that $A \cap B \neq \emptyset$ which implies that $A = B$ so that f is injective.
2. $\mathcal{O} = \{Q_n : n \in N\}$.

Let R be the countable set,

$$R := \mathbb{Q} \cap [0,1).$$

We now claim that

$$N^r \cap N^s = \emptyset \text{ if } r \neq s \text{ and} \tag{17.5}$$

$$[0,1) = \cup_{r \in R} N^r. \tag{17.6}$$

Indeed, if $x \in N^r \cap N^s \neq \emptyset$ then $x = r + n \bmod 1$ and $x = s + n' \bmod 1$, then $n - n' \in \mathbb{Q}$, i.e. $Q_n = Q_{n'}$. That is to say, $n = f(Q_n) = f(Q_{n'}) = n'$ and hence that $s = r \bmod 1$, but $s, r \in [0,1)$ implies that $s = r$. Furthermore, if $x \in [0,1)$ and $n := f(Q_x)$, then $x - n = r \in \mathbb{Q}$ and $x \in N^{r \bmod 1}$. Now that we have constructed N, we are ready for the contradiction. By Equations (17.4–17.6) we find

[1] We have used the Axiom of choice here, i.e. $\prod_{A \in \mathcal{F}} (A \cap [0, 1/3]) \neq \emptyset$

$$1 = \mu([0,1)) = \sum_{r \in R} \mu(N^r) = \sum_{r \in R} \mu(N)$$
$$= \begin{cases} \infty \text{ if } \mu(N) > 0 \\ 0 \ \text{ if } \mu(N) = 0 \end{cases}.$$

which is certainly inconsistent. Incidentally we have just produced an example of so called "non – measurable" set. ■

Because of Theorem 17.5, it is necessary to modify Definition 17.1. Theorem 17.5 points out that we will have to give up the idea of trying to measure all subsets of $\mathbb{R}$ but only measure some sub-collections of "measurable" sets. This leads us to the notion of σ – algebra discussed in the next chapter. Our revised notion of a measure will appear in Definition 19.1 of Chapter 19 below.

18

Measurability

18.1 Algebras and σ – Algebras

Definition 18.1. *A collection of subsets $\mathcal{A}$ of a set X is an **algebra** if*

1. $\emptyset, X \in \mathcal{A}$
2. $A \in \mathcal{A}$ *implies that* $A^c \in \mathcal{A}$
3. $\mathcal{A}$ *is closed under finite unions, i.e. if* $A_1, \dots, A_n \in \mathcal{A}$ *then* $A_1 \cup \dots \cup A_n \in \mathcal{A}$.

   In view of conditions 1. and 2., 3. is equivalent to

3′. $\mathcal{A}$ *is closed under finite intersections.*

Definition 18.2. *A collection of subsets $\mathcal{M}$ of X is a σ – **algebra** (or sometimes called a σ – field) if $\mathcal{M}$ is an algebra which also closed under countable unions, i.e. if $\{A_i\}_{i=1}^{\infty} \subset \mathcal{M}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{M}$. (Notice that since $\mathcal{M}$ is also closed under taking complements, $\mathcal{M}$ is also closed under taking countable intersections.) A pair $(X, \mathcal{M})$, where X is a set and $\mathcal{M}$ is a σ – algebra on X, is called a **measurable space.***

The reader should compare these definitions with that of a topology in Definition 10.1. Recall that the elements of a topology are called open sets. Analogously, elements of and algebra $\mathcal{A}$ or a σ – algebra $\mathcal{M}$ will be called **measurable** sets.

Example 18.3. Here are some examples of algebras.

1. $\mathcal{M} = 2^X$, then $\mathcal{M}$ is a topology, an algebra and a σ – algebra.
2. Let $X = \{1, 2, 3\}$, then $\tau = \{\emptyset, X, \{2, 3\}\}$ is a topology on X which is not an algebra.
3. $\tau = \mathcal{A} = \{\{1\}, \{2, 3\}, \emptyset, X\}$ is a topology, an algebra, and a σ – algebra on X. The sets X, $\{1\}$, $\{2, 3\}$, $\emptyset$ are open and closed. The sets $\{1, 2\}$ and $\{1, 3\}$ are neither open nor **closed** and are not measurable.

The reader should compare this example with Example 10.3.

Proposition 18.4. *Let $\mathcal{E}$ be any collection of subsets of X. Then there exists a unique smallest algebra $\mathcal{A}(\mathcal{E})$ and σ – algebra $\sigma(\mathcal{E})$ which contains $\mathcal{E}$.*

Proof. The proof is the same as the analogous Proposition 10.6 for topologies, i.e.

$$\mathcal{A}(\mathcal{E}) := \bigcap \{\mathcal{A} : \mathcal{A} \text{ is an algebra such that } \mathcal{E} \subset \mathcal{A}\}$$

and

$$\sigma(\mathcal{E}) := \bigcap \{\mathcal{M} : \mathcal{M} \text{ is a } \sigma - \text{algebra such that } \mathcal{E} \subset \mathcal{M}\}.$$

■

Example 18.5. Suppose $X = \{1, 2, 3\}$ and $\mathcal{E} = \{\emptyset, X, \{1, 2\}, \{1, 3\}\}$, see Figure 18.1.

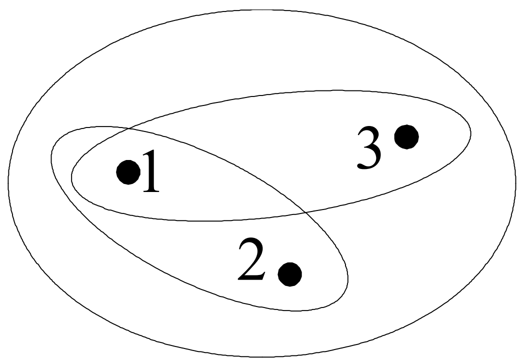


Fig. 18.1. A collection of subsets.

Then

$$\tau(\mathcal{E}) = \{\emptyset, X, \{1\}, \{1, 2\}, \{1, 3\}\}$$
$$\mathcal{A}(\mathcal{E}) = \sigma(\mathcal{E}) = 2^X.$$

The next proposition is the analogue to Proposition 10.7 for topologies and enables us to give and explicit descriptions of $\mathcal{A}(\mathcal{E})$. On the other hand it should be noted that $\sigma(\mathcal{E})$ typically does not admit a simple concrete description.

Proposition 18.6. *Let X be a set and $\mathcal{E} \subset 2^X$. Let $\mathcal{E}^c := \{A^c : A \in \mathcal{E}\}$ and $\mathcal{E}_c := \mathcal{E} \cup \{X, \emptyset\} \cup \mathcal{E}^c$ Then*

$$\mathcal{A}(\mathcal{E}) := \{\text{finite unions of finite intersections of elements from } \mathcal{E}_c\}. \quad (18.1)$$

Proof. Let $\mathcal{A}$ denote the right member of Eq. (18.1). From the definition of an algebra, it is clear that $\mathcal{E} \subset \mathcal{A} \subset \mathcal{A}(\mathcal{E})$. Hence to finish that proof it suffices to show $\mathcal{A}$ is an algebra. The proof of these assertions are routine except for possibly showing that $\mathcal{A}$ is closed under complementation. To check $\mathcal{A}$ is closed under complementation, let $Z \in \mathcal{A}$ be expressed as

$$Z = \bigcup_{i=1}^{N} \bigcap_{j=1}^{K} A_{ij}$$

where $A_{ij} \in \mathcal{E}_c$. Therefore, writing $B_{ij} = A_{ij}^c \in \mathcal{E}_c$, we find that

$$Z^c = \bigcap_{i=1}^{N} \bigcup_{j=1}^{K} B_{ij} = \bigcup_{j_1,\ldots,j_N=1}^{K} (B_{1j_1} \cap B_{2j_2} \cap \cdots \cap B_{Nj_N}) \in \mathcal{A}$$

wherein we have used the fact that $B_{1j_1} \cap B_{2j_2} \cap \cdots \cap B_{Nj_N}$ is a finite intersection of sets from $\mathcal{E}_c$. ∎

Remark 18.7. One might think that in general $\sigma(\mathcal{E})$ may be described as the countable unions of countable intersections of sets in $\mathcal{E}^c$. However this is in general **false,** since if

$$Z = \bigcup_{i=1}^{\infty} \bigcap_{j=1}^{\infty} A_{ij}$$

with $A_{ij} \in \mathcal{E}_c$, then

$$Z^c = \bigcup_{j_1=1, j_2=1, \ldots j_N=1, \ldots}^{\infty} \left(\bigcap_{\ell=1}^{\infty} A_{\ell, j_\ell}^c \right)$$

which is now an **uncountable** union. Thus the above description is not correct. In general it is complicated to explicitly describe $\sigma(\mathcal{E})$, see Proposition 1.23 on page 39 of Folland for details. Also see Proposition 18.13 below.

Exercise 18.1. Let τ be a topology on a set X and $\mathcal{A} = \mathcal{A}(\tau)$ be the algebra generated by τ. Show $\mathcal{A}$ is the collection of subsets of X which may be written as finite union of sets of the form $F \cap V$ where F is closed and V is open.

The following notion will be useful in the sequel and plays an analogous role for algebras as a base (Definition 10.8) does for a topology.

Definition 18.8. *A set $\mathcal{E} \subset 2^X$ is said to be an **elementary family or elementary class** provided that*

- $\emptyset \in \mathcal{E}$
- *$\mathcal{E}$ is closed under finite intersections*
- *if $E \in \mathcal{E}$, then E^c is a finite disjoint union of sets from $\mathcal{E}$. (In particular $X = \emptyset^c$ is a finite disjoint union of elements from $\mathcal{E}$.)*

Example 18.9. Let $X = \mathbb{R}$, then

$$\begin{aligned}\mathcal{E} &:= \left\{(a,b] \cap \mathbb{R} : a,b \in \bar{\mathbb{R}}\right\} \\ &= \{(a,b] : a \in [-\infty,\infty) \text{ and } a < b < \infty\} \cup \{\emptyset, \mathbb{R}\}\end{aligned}$$

is an elementary family.

Exercise 18.2. Let $\mathcal{A} \subset 2^X$ and $\mathcal{B} \subset 2^Y$ be elementary families. Show the collection

$$\mathcal{E} = \mathcal{A} \times \mathcal{B} = \{A \times B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$$

is also an elementary family.

Proposition 18.10. *Suppose $\mathcal{E} \subset 2^X$ is an elementary family, then $\mathcal{A} = \mathcal{A}(\mathcal{E})$ consists of sets which may be written as finite disjoint unions of sets from $\mathcal{E}$.*

Proof. This could be proved making use of Proposition 18.6. However it is easier to give a direct proof. Let $\mathcal{A}$ denote the collection of sets which may be written as finite disjoint unions of sets from $\mathcal{E}$. Clearly $\mathcal{E} \subset \mathcal{A} \subset \mathcal{A}(\mathcal{E})$ so it suffices to show $\mathcal{A}$ is an algebra since $\mathcal{A}(\mathcal{E})$ is the smallest algebra containing $\mathcal{E}$. By the properties of $\mathcal{E}$, we know that $\emptyset, X \in \mathcal{A}$. Now suppose that $A_i = \coprod_{F \in \Lambda_i} F \in \mathcal{A}$ where, for $i = 1, 2, \ldots, n$, Λ_i is a finite collection of disjoint sets from $\mathcal{E}$. Then

$$\bigcap_{i=1}^n A_i = \bigcap_{i=1}^n \left(\coprod_{F \in \Lambda_i} F\right) = \bigcup_{(F_1,,\ldots,F_n) \in \Lambda_1 \times \cdots \times \Lambda_n} (F_1 \cap F_2 \cap \cdots \cap F_n)$$

and this is a disjoint (you check) union of elements from $\mathcal{E}$. Therefore $\mathcal{A}$ is closed under finite intersections. Similarly, if $A = \coprod_{F \in \Lambda} F$ with Λ being a finite collection of disjoint sets from $\mathcal{E}$, then $A^c = \bigcap_{F \in \Lambda} F^c$. Since by assumption $F^c \in \mathcal{A}$ for $F \in \Lambda \subset \mathcal{E}$ and $\mathcal{A}$ is closed under finite intersections, it follows that $A^c \in \mathcal{A}$. ■

Definition 18.11. *Let X be a set. We say that a family of sets $\mathcal{F} \subset 2^X$ is a **partition** of X if distinct members of $\mathcal{F}$ are disjoint and if X is the union of the sets in $\mathcal{F}$.*

Example 18.12. Let X be a set and $\mathcal{E} = \{A_1, \ldots, A_n\}$ where $A_1, \ldots, A_n$ is a partition of X. In this case

$$\mathcal{A}(\mathcal{E}) = \sigma(\mathcal{E}) = \tau(\mathcal{E}) = \{\cup_{i \in \Lambda} A_i : \Lambda \subset \{1, 2, \ldots, n\}\}$$

where $\cup_{i \in \Lambda} A_i := \emptyset$ when $\Lambda = \emptyset$. Notice that

$$\#\left(\mathcal{A}(\mathcal{E})\right) = \#(2^{\{1,2,\ldots,n\}}) = 2^n.$$

Proposition 18.13. *Suppose that $\mathcal{M} \subset 2^X$ is a σ – algebra and $\mathcal{M}$ is at most a countable set. Then there exists a unique **finite** partition $\mathcal{F}$ of X such that $\mathcal{F} \subset \mathcal{M}$ and every element $B \in \mathcal{M}$ is of the form*

$$B = \cup \{A \in \mathcal{F} : A \subset B\}. \tag{18.2}$$

In particular $\mathcal{M}$ is actually a finite set and $\#(\mathcal{M}) = 2^n$ for some $n \in \mathbb{N}$.

Proof. For each $x \in X$ let

$$A_x = \cap \{A \in \mathcal{M} : x \in A\} \in \mathcal{M},$$

wherein we have used $\mathcal{M}$ is a countable σ – algebra to insure $A_x \in \mathcal{M}$. Hence A_x is the smallest set in $\mathcal{M}$ which contains x. Let $C = A_x \cap A_y$. If $x \notin C$ then $A_x \setminus C \subset A_x$ is an element of $\mathcal{M}$ which contains x and since A_x is the smallest member of $\mathcal{M}$ containing x, we must have that $C = \emptyset$. Similarly if $y \notin C$ then $C = \emptyset$. Therefore if $C \neq \emptyset$, then $x, y \in A_x \cap A_y \in \mathcal{M}$ and $A_x \cap A_y \subset A_x$ and $A_x \cap A_y \subset A_y$ from which it follows that $A_x = A_x \cap A_y = A_y$. This shows that $\mathcal{F} = \{A_x : x \in X\} \subset \mathcal{M}$ is a (necessarily countable) partition of X for which Eq. (18.2) holds for all $B \in \mathcal{M}$. Enumerate the elements of $\mathcal{F}$ as $\mathcal{F} = \{P_n\}_{n=1}^N$ where $N \in \mathbb{N}$ or $N = \infty$. If $N = \infty$, then the correspondence

$$a \in \{0,1\}^{\mathbb{N}} \to A_a = \cup \{P_n : a_n = 1\} \in \mathcal{M}$$

is bijective and therefore, by Lemma 2.6, $\mathcal{M}$ is uncountable. Thus any countable σ – algebra is necessarily finite. This finishes the proof modulo the uniqueness assertion which is left as an exercise to the reader. ■

Example 18.14. Let $X = \mathbb{R}$ and

$$\mathcal{E} = \{(a, \infty) : a \in \mathbb{R}\} \cup \{\mathbb{R}, \emptyset\} = \{(a, \infty) \cap \mathbb{R} : a \in \bar{\mathbb{R}}\} \subset 2^{\mathbb{R}}.$$

Notice that $\mathcal{E}_f = \mathcal{E}$ and that $\mathcal{E}$ is closed under unions, which shows that $\tau(\mathcal{E}) = \mathcal{E}$, i.e. $\mathcal{E}$ is already a topology. Since $(a, \infty)^c = (-\infty, a]$ we find that $\mathcal{E}_c = \{(a, \infty), (-\infty, a], -\infty \leq a < \infty\} \cup \{\mathbb{R}, \emptyset\}$. Noting that

$$(a, \infty) \cap (-\infty, b] = (a, b]$$

it follows that $\mathcal{A}(\mathcal{E}) = \mathcal{A}(\tilde{\mathcal{E}})$ where

$$\tilde{\mathcal{E}} := \left\{(a, b] \cap \mathbb{R} : a, b \in \bar{\mathbb{R}}\right\}.$$

Since $\tilde{\mathcal{E}}$ is an elementary family of subsets of $\mathbb{R}$, Proposition 18.10 implies $\mathcal{A}(\mathcal{E})$ may be described as being those sets which are finite disjoint unions of sets from $\tilde{\mathcal{E}}$. The σ – algebra, $\sigma(\mathcal{E})$, generated by $\mathcal{E}$ is **very complicated**. Here are some sets in $\sigma(\mathcal{E})$ – most of which are not in $\mathcal{A}(\mathcal{E})$.

(a) $(a, b) = \bigcup_{n=1}^{\infty} (a, b - \frac{1}{n}] \in \sigma(\mathcal{E})$.

(b) All of the standard open subsets of $\mathbb{R}$ are in $\sigma(\mathcal{E})$.
(c) $\{x\} = \bigcap_n \left(x - \frac{1}{n}, x\right] \in \sigma(\mathcal{E})$
(d) $[a, b] = \{a\} \cup (a, b] \in \sigma(\mathcal{E})$
(e) Any countable subset of $\mathbb{R}$ is in $\sigma(\mathcal{E})$.

Remark 18.15. In the above example, one may replace $\mathcal{E}$ by $\mathcal{E} = \{(a, \infty) : a \in \mathbb{Q}\} \cup \{\mathbb{R}, \emptyset\}$, in which case $\mathcal{A}(\mathcal{E})$ may be described as being those sets which are finite disjoint unions of sets from the following list

$$\{(a, \infty), (-\infty, a], (a, b] : a, b \in \mathbb{Q}\} \cup \{\emptyset, \mathbb{R}\}.$$

This shows that $\mathcal{A}(\mathcal{E})$ is a countable set – a useful fact which will be needed later.

Notation 18.16 *For a general topological space (X, τ), the **Borel** σ – **algebra** is the σ – algebra $\mathcal{B}_X := \sigma(\tau)$ on X. In particular if $X = \mathbb{R}^n$, $\mathcal{B}_{\mathbb{R}^n}$ will be used to denote the Borel σ – algebra on $\mathbb{R}^n$ when $\mathbb{R}^n$ is equipped with its standard Euclidean topology.*

Exercise 18.3. Verify the σ – algebra, $\mathcal{B}_{\mathbb{R}}$, is generated by any of the following collection of sets:

$$1.\ \{(a, \infty) : a \in \mathbb{R}\},\ 2.\ \{(a, \infty) : a \in \mathbb{Q}\} \text{ or } 3.\ \{[a, \infty) : a \in \mathbb{Q}\}.$$

Proposition 18.17. *If τ is a second countable topology on X and $\mathcal{E}$ is a countable collection of subsets of X such that $\tau = \tau(\mathcal{E})$, then $\mathcal{B}_X := \sigma(\tau) = \sigma(\mathcal{E})$, i.e. $\sigma(\tau(\mathcal{E})) = \sigma(\mathcal{E})$.*

Proof. Let $\mathcal{E}_f$ denote the collection of subsets of X which are finite intersection of elements from $\mathcal{E}$ along with X and $\emptyset$. Notice that $\mathcal{E}_f$ is still countable (you prove). A set Z is in $\tau(\mathcal{E})$ iff Z is an arbitrary union of sets from $\mathcal{E}_f$. Therefore $Z = \bigcup_{A \in \mathcal{F}} A$ for some subset $\mathcal{F} \subset \mathcal{E}_f$ which is necessarily countable. Since $\mathcal{E}_f \subset \sigma(\mathcal{E})$ and $\sigma(\mathcal{E})$ is closed under countable unions it follows that $Z \in \sigma(\mathcal{E})$ and hence that $\tau(\mathcal{E}) \subset \sigma(\mathcal{E})$. Lastly, since $\mathcal{E} \subset \tau(\mathcal{E}) \subset \sigma(\mathcal{E})$, $\sigma(\mathcal{E}) \subset \sigma(\tau(\mathcal{E})) \subset \sigma(\mathcal{E})$. ■

18.2 Measurable Functions

Our notion of a "measurable" function will be analogous to that for a continuous function. For motivational purposes, suppose $(X, \mathcal{M}, \mu)$ is a measure space and $f : X \to \mathbb{R}_+$. Roughly speaking, in the next Chapter we are going to define $\int_X f d\mu$ as a certain limit of sums of the form,

$$\sum_{0 < a_1 < a_2 < a_3 < \dots}^{\infty} a_i \mu(f^{-1}(a_i, a_{i+1}]).$$

For this to make sense we will need to require $f^{-1}((a,b]) \in \mathcal{M}$ for all $a < b$. Because of Lemma 18.22 below, this last condition is equivalent to the condition $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{M}$.

Definition 18.18. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{F})$ be measurable spaces. A function $f : X \to Y$ is* ***measurable*** *if $f^{-1}(\mathcal{F}) \subset \mathcal{M}$. We will also say that f is $\mathcal{M}/\mathcal{F}$ – measurable or $(\mathcal{M}, \mathcal{F})$ – measurable.*

Example 18.19 (Characteristic Functions). Let $(X, \mathcal{M})$ be a measurable space and $A \subset X$. We define the **characteristic function** $1_A : X \to \mathbb{R}$ by

$$1_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A. \end{cases}$$

If $A \in \mathcal{M}$, then 1_A is $(\mathcal{M}, 2^{\mathbb{R}})$ – measurable because $1_A^{-1}(W)$ is either $\emptyset$, X, A or A^c for any $W \subset \mathbb{R}$. Conversely, if $\mathcal{F}$ is any σ – algebra on $\mathbb{R}$ containing a set $W \subset \mathbb{R}$ such that $1 \in W$ and $0 \in W^c$, then $A \in \mathcal{M}$ if 1_A is $(\mathcal{M}, \mathcal{F})$ – measurable. This is because $A = 1_A^{-1}(W) \in \mathcal{M}$.

Exercise 18.4. Suppose $f : X \to Y$ is a function, $\mathcal{F} \subset 2^Y$ and $\mathcal{M} \subset 2^X$. Show $f^{-1}\mathcal{F}$ and $f_*\mathcal{M}$ (see Notation 2.7) are algebras (σ – algebras) provided $\mathcal{F}$ and $\mathcal{M}$ are algebras (σ – algebras).

Remark 18.20. Let $f : X \to Y$ be a function. Given a σ – algebra $\mathcal{F} \subset 2^Y$, the σ – algebra $\mathcal{M} := f^{-1}(\mathcal{F})$ is the smallest σ – algebra on X such that f is $(\mathcal{M}, \mathcal{F})$ - measurable . Similarly, if $\mathcal{M}$ is a σ - algebra on X then $\mathcal{F} = f_*\mathcal{M}$ is the largest σ – algebra on Y such that f is $(\mathcal{M}, \mathcal{F})$ - measurable .

Recall from Definition 2.8 that for $\mathcal{E} \subset 2^X$ and $A \subset X$ that

$$\mathcal{E}_A = i_A^{-1}(\mathcal{E}) = \{A \cap E : E \in \mathcal{E}\}$$

where $i_A : A \to X$ is the inclusion map. Because of Exercise 10.3, when $\mathcal{E} = \mathcal{M}$ is an algebra (σ – algebra), $\mathcal{M}_A$ is an algebra (σ – algebra) on A and we call $\mathcal{M}_A$ the relative or induced algebra (σ – algebra) on A.

The next two Lemmas are direct analogues of their topological counter parts in Lemmas 10.13 and 10.14. For completeness, the proofs will be given even though they are same as those for Lemmas 10.13 and 10.14.

Lemma 18.21. *Suppose that $(X, \mathcal{M})$, $(Y, \mathcal{F})$ and $(Z, \mathcal{G})$ are measurable spaces. If $f : (X, \mathcal{M}) \to (Y, \mathcal{F})$ and $g : (Y, \mathcal{F}) \to (Z, \mathcal{G})$ are measurable functions then $g \circ f : (X, \mathcal{M}) \to (Z, \mathcal{G})$ is measurable as well.*

Proof. By assumption $g^{-1}(\mathcal{G}) \subset \mathcal{F}$ and $f^{-1}(\mathcal{F}) \subset \mathcal{M}$ so that

$$(g \circ f)^{-1}(\mathcal{G}) = f^{-1}\left(g^{-1}(\mathcal{G})\right) \subset f^{-1}(\mathcal{F}) \subset \mathcal{M}.$$

■

Lemma 18.22. *Suppose that $f : X \to Y$ is a function and $\mathcal{E} \subset 2^Y$ and $A \subset Y$ then*

$$\sigma\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\sigma(\mathcal{E})) \text{ and} \tag{18.3}$$

$$(\sigma(\mathcal{E}))_A = \sigma(\mathcal{E}_A\ ). \tag{18.4}$$

(Similar assertion hold with $\sigma(\cdot)$ being replaced by $\mathcal{A}(\cdot)$.) Moreover, if $\mathcal{F} = \sigma(\mathcal{E})$ and $\mathcal{M}$ is a σ – algebra on X, then f is $(\mathcal{M}, \mathcal{F})$ – measurable iff $f^{-1}(\mathcal{E}) \subset \mathcal{M}$.

Proof. By Exercise 18.4, $f^{-1}(\sigma(\mathcal{E}))$ is a σ – algebra and since $\mathcal{E} \subset \mathcal{F}$, $f^{-1}(\mathcal{E}) \subset f^{-1}(\sigma(\mathcal{E}))$. It now follows that $\sigma\ (f^{-1}(\mathcal{E})) \subset f^{-1}(\sigma\ (\mathcal{E}))$. For the reverse inclusion, notice that

$$f_*\sigma\left(f^{-1}(\mathcal{E})\right) = \left\{B \subset Y : f^{-1}(B) \in \sigma\left(f^{-1}(\mathcal{E})\right)\right\}$$

is a σ – algebra which contains $\mathcal{E}$ and thus $\sigma(\mathcal{E}) \subset f_*\sigma\ \left(f^{-1}(\mathcal{E})\right)$. Hence if $B \in \sigma(\mathcal{E})$ we know that $f^{-1}(B) \in \sigma\left(f^{-1}(\mathcal{E})\right)$, i.e. $f^{-1}(\sigma(\mathcal{E})) \subset \sigma\left(f^{-1}(\mathcal{E})\right)$ and Eq. (18.3) has been proved. Applying Eq. (18.3) with $X = A$ and $f = i_A$ being the inclusion map implies

$$(\sigma(\mathcal{E}))_A = i_A^{-1}(\sigma(\mathcal{E})) = \sigma(i_A^{-1}(\mathcal{E})) = \sigma(\mathcal{E}_A).$$

Lastly if $f^{-1}\mathcal{E} \subset \mathcal{M}$, then $f^{-1}\sigma(\mathcal{E}) = \sigma\left(f^{-1}\mathcal{E}\right) \subset \mathcal{M}$ which shows f is $(\mathcal{M}, \mathcal{F})$ – measurable. ∎

Corollary 18.23. *Suppose that $(X, \mathcal{M})$ is a measurable space. Then the following conditions on a function $f : X \to \mathbb{R}$ are equivalent:*

1. *f is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ – measurable,*
2. *$f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$,*
3. *$f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{Q}$,*
4. *$f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$.*

Proof. An exercise in using Lemma 18.22 and is the content of Exercise 18.8. ∎

Here is yet another way to generate σ – algebras. (Compare with the analogous topological Definition 10.20.)

Definition 18.24 (σ – Algebras Generated by Functions). *Let X be a set and suppose there is a collection of measurable spaces $\{(Y_\alpha, \mathcal{F}_\alpha) : \alpha \in A\}$ and functions $f_\alpha : X \to Y_\alpha$ for all $\alpha \in A$. Let $\sigma(f_\alpha : \alpha \in A)$ denote the smallest σ – algebra on X such that each f_α is measurable, i.e.*

$$\sigma(f_\alpha : \alpha \in A) = \sigma(\cup_\alpha f_\alpha^{-1}(\mathcal{F}_\alpha)).$$

Proposition 18.25. *Assuming the notation in Definition 18.24 and additionally let $(Z, \mathcal{M})$ be a measurable space and $g : Z \to X$ be a function. Then g is $(\mathcal{M}, \sigma(f_\alpha : \alpha \in A))$ – measurable iff $f_\alpha \circ g$ is $(\mathcal{M}, \mathcal{F}_\alpha)$–measurable for all $\alpha \in A$.*

Proof. This proof is essentially the same as the proof of the topological analogue in Proposition 10.21. ($\Rightarrow$) If g is $(\mathcal{M}, \sigma(f_\alpha : \alpha \in A))$ – measurable, then the composition $f_\alpha \circ g$ is $(\mathcal{M}, \mathcal{F}_\alpha)$ – measurable by Lemma 18.21. ($\Leftarrow$) Let

$$\mathcal{G} = \sigma(f_\alpha : \alpha \in A) = \sigma\left(\cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{F}_\alpha)\right).$$

If $f_\alpha \circ g$ is $(\mathcal{M}, \mathcal{F}_\alpha)$ – measurable for all α, then

$$g^{-1} f_\alpha^{-1}(\mathcal{F}_\alpha) \subset \mathcal{M}\, \forall\, \alpha \in A$$

and therefore

$$g^{-1}\left(\cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{F}_\alpha)\right) = \cup_{\alpha\in A} g^{-1} f_\alpha^{-1}(\mathcal{F}_\alpha) \subset \mathcal{M}.$$

Hence

$$g^{-1}(\mathcal{G}) = g^{-1}\left(\sigma\left(\cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{F}_\alpha)\right)\right) = \sigma(g^{-1}\left(\cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{F}_\alpha)\right) \subset \mathcal{M}$$

which shows that g is $(\mathcal{M}, \mathcal{G})$ – measurable. ∎

Definition 18.26. *A function $f : X \to Y$ between two topological spaces is* ***Borel measurable*** *if $f^{-1}(\mathcal{B}_Y) \subset \mathcal{B}_X$.*

Proposition 18.27. *Let X and Y be two topological spaces and $f : X \to Y$ be a continuous function. Then f is Borel measurable.*

Proof. Using Lemma 18.22 and $\mathcal{B}_Y = \sigma(\tau_Y)$,

$$f^{-1}(\mathcal{B}_Y) = f^{-1}(\sigma(\tau_Y)) = \sigma(f^{-1}(\tau_Y)) \subset \sigma(\tau_X) = \mathcal{B}_X.$$

∎

Definition 18.28. *Given measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{F})$ and a subset $A \subset X$. We say a function $f : A \to Y$ is measurable iff f is $\mathcal{M}_A/\mathcal{F}$ – measurable.*

Proposition 18.29 (Localizing Measurability). *Let $(X, \mathcal{M})$ and $(Y, \mathcal{F})$ be measurable spaces and $f : X \to Y$ be a function.*

1. *If f is measurable and $A \subset X$ then $f|_A : A \to Y$ is measurable.*
2. *Suppose there exist $A_n \in \mathcal{M}$ such that $X = \cup_{n=1}^\infty A_n$ and $f|A_n$ is $\mathcal{M}_{A_n}$ measurable for all n, then f is $\mathcal{M}$ – measurable.*

Proof. As the reader will notice, the proof given below is essentially identical to the proof of Proposition 10.19 which is the topological analogue of this proposition. 1. If $f : X \to Y$ is measurable, $f^{-1}(B) \in \mathcal{M}$ for all $B \in \mathcal{F}$ and therefore

$$f|_A^{-1}(B) = A \cap f^{-1}(B) \in \mathcal{M}_A \text{ for all } B \in \mathcal{F}.$$

2. If $B \in \mathcal{F}$, then

$$f^{-1}(B) = \cup_{n=1}^{\infty}\left(f^{-1}(B) \cap A_n\right) = \cup_{n=1}^{\infty} f|_{A_n}^{-1}(B).$$

Since each $A_n \in \mathcal{M}$, $\mathcal{M}_{A_n} \subset \mathcal{M}$ and so the previous displayed equation shows $f^{-1}(B) \in \mathcal{M}$. ■

Proposition 18.30. *If $(X, \mathcal{M})$ is a measurable space, then*

$$f = (f_1, f_2, \ldots, f_n) : X \to \mathbb{R}^n$$

is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}^n})$ – measurable iff $f_i : X \to \mathbb{R}$ is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ – measurable for each i. In particular, a function $f : X \to \mathbb{C}$ is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ – measurable iff $\operatorname{Re} f$ *and* $\operatorname{Im} f$ *are $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ – measurable.*

Proof. This is formally a consequence of Corollary 18.65 and Proposition 18.60 below. Nevertheless it is instructive to give a direct proof now. Let $\tau = \tau_{\mathbb{R}^n}$ denote the usual topology on $\mathbb{R}^n$ and $\pi_i : \mathbb{R}^n \to \mathbb{R}$ be projection onto the i^{th} – factor. Since π_i is continuous, π_i is $\mathcal{B}_{\mathbb{R}^n}/\mathcal{B}_{\mathbb{R}}$ – measurable and therefore if $f : X \to \mathbb{R}^n$ is measurable then so is $f_i = \pi_i \circ f$. Now suppose $f_i : X \to \mathbb{R}$ is measurable for all $i = 1, 2, \ldots, n$. Let

$$\mathcal{E} := \{(a, b) : a, b \in \mathbb{Q}^n \ni a < b\},$$

where, for $a, b \in \mathbb{R}^n$, we write $a < b$ iff $a_i < b_i$ for $i = 1, 2, \ldots, n$ and let

$$(a, b) = (a_1, b_1) \times \cdots \times (a_n, b_n).$$

Since $\mathcal{E} \subset \tau$ and every element $V \in \tau$ may be written as a (necessarily) countable union of elements from $\mathcal{E}$, we have $\sigma(\mathcal{E}) \subset \mathcal{B}_{\mathbb{R}^n} = \sigma(\tau) \subset \sigma(\mathcal{E})$, i.e. $\sigma(\mathcal{E}) = \mathcal{B}_{\mathbb{R}^n}$. (This part of the proof is essentially a direct proof of Corollary 18.65 below.) Because

$$f^{-1}((a, b)) = f_1^{-1}((a_1, b_1)) \cap f_2^{-1}((a_2, b_2)) \cap \cdots \cap f_n^{-1}((a_n, b_n)) \in \mathcal{M}$$

for all $a, b \in \mathbb{Q}$ with $a < b$, it follows that $f^{-1}\mathcal{E} \subset \mathcal{M}$ and therefore

$$f^{-1}\mathcal{B}_{\mathbb{R}^n} = f^{-1}\sigma(\mathcal{E}) = \sigma\left(f^{-1}\mathcal{E}\right) \subset \mathcal{M}.$$

■

Corollary 18.31. *Let $(X, \mathcal{M})$ be a measurable space and $f, g : X \to \mathbb{C}$ be $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ – measurable functions. Then $f \pm g$ and $f \cdot g$ are also $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ – measurable.*

Proof. Define $F : X \to \mathbb{C} \times \mathbb{C}$, $A_{\pm} : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and $M : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}$ by $F(x) = (f(x), g(x))$, $A_{\pm}(w, z) = w \pm z$ and $M(w, z) = wz$. Then $A_{\pm}$ and M are continuous and hence $(\mathcal{B}_{\mathbb{C}^2}, \mathcal{B}_{\mathbb{C}})$ – measurable. Also F is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}} \otimes \mathcal{B}_{\mathbb{C}}) = (\mathcal{M}, \mathcal{B}_{\mathbb{C}^2})$ – measurable since $\pi_1 \circ F = f$ and $\pi_2 \circ F = g$ are $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ – measurable. Therefore $A_{\pm} \circ F = f \pm g$ and $M \circ F = f \cdot g$, being the composition of measurable functions, are also measurable. ■

Lemma 18.32. *Let $\alpha \in \mathbb{C}$, $(X, \mathcal{M})$ be a measurable space and $f : X \to \mathbb{C}$ be a $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ – measurable function. Then*

$$F(x) := \begin{cases} \frac{1}{f(x)} & \text{if} \quad f(x) \neq 0 \\ \alpha & \text{if} \quad f(x) = 0 \end{cases}$$

is measurable.

Proof. Define $i : \mathbb{C} \to \mathbb{C}$ by

$$i(z) = \begin{cases} \frac{1}{z} \text{ if} & z \neq 0 \\ 0 \text{ if} & z = 0. \end{cases}$$

For any open set $V \subset \mathbb{C}$ we have

$$i^{-1}(V) = i^{-1}(V \setminus \{0\}) \cup i^{-1}(V \cap \{0\})$$

Because i is continuous except at $z = 0$, $i^{-1}(V \setminus \{0\})$ is an open set and hence in $\mathcal{B}_{\mathbb{C}}$. Moreover, $i^{-1}(V \cap \{0\}) \in \mathcal{B}_{\mathbb{C}}$ since $i^{-1}(V \cap \{0\})$ is either the empty set or the one point set $\{0\}$. Therefore $i^{-1}(\tau_{\mathbb{C}}) \subset \mathcal{B}_{\mathbb{C}}$ and hence $i^{-1}(\mathcal{B}_{\mathbb{C}}) = i^{-1}(\sigma(\tau_{\mathbb{C}})) = \sigma(i^{-1}(\tau_{\mathbb{C}})) \subset \mathcal{B}_{\mathbb{C}}$ which shows that i is Borel measurable. Since $F = i \circ f$ is the composition of measurable functions, F is also measurable. ■

We will often deal with functions $f : X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. When talking about measurability in this context we will refer to the σ – algebra on $\bar{\mathbb{R}}$ defined by

$$\mathcal{B}_{\bar{\mathbb{R}}} := \sigma\left(\{[a, \infty] : a \in \mathbb{R}\}\right). \tag{18.5}$$

Proposition 18.33 (The Structure of $\mathcal{B}_{\bar{\mathbb{R}}}$). *Let $\mathcal{B}_{\mathbb{R}}$ and $\mathcal{B}_{\bar{\mathbb{R}}}$ be as above, then*

$$\mathcal{B}_{\bar{\mathbb{R}}} = \{A \subset \bar{\mathbb{R}} : A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\}. \tag{18.6}$$

In particular $\{\infty\}, \{-\infty\} \in \mathcal{B}_{\bar{\mathbb{R}}}$ and $\mathcal{B}_{\mathbb{R}} \subset \mathcal{B}_{\bar{\mathbb{R}}}$.

Proof. Let us first observe that

$$\begin{aligned} \{-\infty\} &= \cap_{n=1}^{\infty} [-\infty, -n) = \cap_{n=1}^{\infty} [-n, \infty]^c \in \mathcal{B}_{\bar{\mathbb{R}}}, \\ \{\infty\} &= \cap_{n=1}^{\infty} [n, \infty] \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ and } \mathbb{R} = \bar{\mathbb{R}} \setminus \{\pm\infty\} \in \mathcal{B}_{\bar{\mathbb{R}}}. \end{aligned}$$

Letting $i : \mathbb{R} \to \bar{\mathbb{R}}$ be the inclusion map,

$$\begin{aligned} i^{-1}\left(\mathcal{B}_{\bar{\mathbb{R}}}\right) &= \sigma\left(i^{-1}\left(\left\{[a, \infty] : a \in \bar{\mathbb{R}}\right\}\right)\right) = \sigma\left(\left\{i^{-1}\left([a, \infty]\right) : a \in \bar{\mathbb{R}}\right\}\right) \\ &= \sigma\left(\left\{[a, \infty] \cap \mathbb{R} : a \in \bar{\mathbb{R}}\right\}\right) = \sigma\left(\{[a, \infty) : a \in \mathbb{R}\}\right) = \mathcal{B}_{\mathbb{R}}. \end{aligned}$$

Thus we have shown

$$\mathcal{B}_{\mathbb{R}} = i^{-1}\left(\mathcal{B}_{\bar{\mathbb{R}}}\right) = \{A \cap \mathbb{R} : A \in \mathcal{B}_{\bar{\mathbb{R}}}\}.$$

This implies:

1. $A \in \mathcal{B}_{\bar{\mathbb{R}}} \Longrightarrow A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ and
2. if $A \subset \bar{\mathbb{R}}$ is such that $A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$ there exists $B \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $A \cap \mathbb{R} = B \cap \mathbb{R}$. Because $A \Delta B \subset \{\pm\infty\}$ and $\{\infty\}, \{-\infty\} \in \mathcal{B}_{\bar{\mathbb{R}}}$ we may conclude that $A \in \mathcal{B}_{\bar{\mathbb{R}}}$ as well.

This proves Eq. (18.6). ■

The proofs of the next two corollaries are left to the reader, see Exercises 18.5 and 18.6.

Corollary 18.34. *Let $(X, \mathcal{M})$ be a measurable space and $f : X \to \bar{\mathbb{R}}$ be a function. Then the following are equivalent*

1. *f is $(\mathcal{M}, \mathcal{B}_{\bar{\mathbb{R}}})$ - measurable,*
2. *$f^{-1}((a, \infty]) \in \mathcal{M}$ for all $a \in \mathbb{R}$,*
3. *$f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$,*
4. *$f^{-1}(\{-\infty\}) \in \mathcal{M}$, $f^{-1}(\{\infty\}) \in \mathcal{M}$ and $f^0 : X \to \mathbb{R}$ defined by*

$$f^0(x) := 1_{\mathbb{R}}(f(x)) = \begin{cases} f(x) & \text{if } \quad f(x) \in \mathbb{R} \\ 0 & \text{if } f(x) \in \{\pm\infty\} \end{cases}$$

is measurable.

Corollary 18.35. *Let $(X, \mathcal{M})$ be a measurable space, $f, g : X \to \bar{\mathbb{R}}$ be functions and define $f \cdot g : X \to \bar{\mathbb{R}}$ and $(f + g) : X \to \bar{\mathbb{R}}$ using the conventions, $0 \cdot \infty = 0$ and $(f + g)(x) = 0$ if $f(x) = \infty$ and $g(x) = -\infty$ or $f(x) = -\infty$ and $g(x) = \infty$. Then $f \cdot g$ and $f + g$ are measurable functions on X if both f and g are measurable.*

Exercise 18.5. Prove Corollary 18.34 noting that the equivalence of items 1. – 3. is a direct analogue of Corollary 18.23. Use Proposition 18.33 to handle item 4.

Exercise 18.6. Prove Corollary 18.35.

Proposition 18.36 (Closure under sups, infs and limits). *Suppose that $(X, \mathcal{M})$ is a measurable space and $f_j : (X, \mathcal{M}) \to \overline{\mathbb{R}}$ for $j \in \mathbb{N}$ is a sequence of $\mathcal{M}/\mathcal{B}_{\overline{\mathbb{R}}}$ – measurable functions. Then*

$$\sup_j f_j, \quad \inf_j f_j, \quad \limsup_{j \to \infty} f_j \text{ and } \liminf_{j \to \infty} f_j$$

are all $\mathcal{M}/\mathcal{B}_{\overline{\mathbb{R}}}$ – measurable functions. (Note that this result is in generally false when $(X, \mathcal{M})$ is a topological space and measurable is replaced by continuous in the statement.)

Proof. Define $g_+(x) := \sup_j f_j(x)$, then

$$\begin{aligned} \{x : g_+(x) \le a\} &= \{x : f_j(x) \le a \ \forall j\} \\ &= \cap_j \{x : f_j(x) \le a\} \in \mathcal{M} \end{aligned}$$

so that g_+ is measurable. Similarly if $g_-(x) = \inf_j f_j(x)$ then

$$\{x : g_-(x) \geq a\} = \cap_j \{x : f_j(x) \geq a\} \in \mathcal{M}.$$

Since

$$\limsup_{j\to\infty} f_j = \inf_n \sup\{f_j : j \geq n\} \text{ and}$$
$$\liminf_{j\to\infty} f_j = \sup_n \inf\{f_j : j \geq n\}$$

we are done by what we have already proved. ∎

Definition 18.37. *Given a function $f : X \to \bar{\mathbb{R}}$ let $f_+(x) := \max\{f(x), 0\}$ and $f_-(x) := \max(-f(x), 0) = -\min(f(x), 0)$. Notice that $f = f_+ - f_-$.*

Corollary 18.38. *Suppose $(X, \mathcal{M})$ is a measurable space and $f : X \to \bar{\mathbb{R}}$ is a function. Then f is measurable iff $f_\pm$ are measurable.*

Proof. If f is measurable, then Proposition 18.36 implies $f_\pm$ are measurable. Conversely if $f_\pm$ are measurable then so is $f = f_+ - f_-$. ∎

18.2.1 More general pointwise limits

Lemma 18.39. *Suppose that $(X, \mathcal{M})$ is a measurable space, (Y, d) is a metric space and $f_j : X \to Y$ is $(\mathcal{M}, \mathcal{B}_Y)$ – measurable for all j. Also assume that for each $x \in X$, $f(x) = \lim_{n\to\infty} f_n(x)$ exists. Then $f : X \to Y$ is also $(\mathcal{M}, \mathcal{B}_Y)$ – measurable.*

Proof. Let $V \in \tau_d$ and $W_m := \{y \in Y : d_{V^c}(y) > 1/m\}$ for $m = 1, 2, \dots$. Then $W_m \in \tau_d$,

$$W_m \subset \bar{W}_m \subset \{y \in Y : d_{V^c}(y) \geq 1/m\} \subset V$$

for all m and $W_m \uparrow V$ as $m \to \infty$. The proof will be completed by verifying the identity,

$$f^{-1}(V) = \cup_{m=1}^\infty \cup_{N=1}^\infty \cap_{n\geq N} f_n^{-1}(W_m) \in \mathcal{M}.$$

If $x \in f^{-1}(V)$ then $f(x) \in V$ and hence $f(x) \in W_m$ for some m. Since $f_n(x) \to f(x)$, $f_n(x) \in W_m$ for almost all n. That is $x \in \cup_{m=1}^\infty \cup_{N=1}^\infty \cap_{n\geq N} f_n^{-1}(W_m)$. Conversely when $x \in \cup_{m=1}^\infty \cup_{N=1}^\infty \cap_{n\geq N} f_n^{-1}(W_m)$ there exists an m such that $f_n(x) \in W_m \subset \bar{W}_m$ for almost all n. Since $f_n(x) \to f(x) \in \bar{W}_m \subset V$, it follows that $x \in f^{-1}(V)$. ∎

Remark 18.40. In the previous Lemma 18.39 it is possible to let (Y, τ) be any topological space which has the "regularity" property that if $V \in \tau$ there exists $W_m \in \tau$ such that $W_m \subset \bar{W}_m \subset V$ and $V = \cup_{m=1}^\infty W_m$. Moreover, some extra condition is necessary on the topology τ in order for Lemma 18.39 to be correct. For example if $Y = \{1, 2, 3\}$ and $\tau = \{Y, \emptyset, \{1, 2\}, \{2, 3\}, \{2\}\}$ as in Example 10.36 and $X = \{a, b\}$ with the trivial σ – algebra. Let $f_j(a) = f_j(b) = 2$ for all j, then f_j is constant and hence measurable. Let $f(a) = 1$ and $f(b) = 2$, then $f_j \to f$ as $j \to \infty$ with f being non-measurable. Notice that the Borel σ – algebra on Y is 2^Y.

18.3 σ – Function Algebras

In this subsection, we are going to relate σ – algebras of subsets of a set X to certain algebras of functions on X. We will begin this endeavor after proving the simple but very useful approximation Theorem 18.42 below.

Definition 18.41. *Let $(X, \mathcal{M})$ be a measurable space. A function $\phi : X \to \mathbb{F}$ ($\mathbb{F}$ denotes either $\mathbb{R}$, $\mathbb{C}$ or $[0,\infty] \subset \bar{\mathbb{R}}$) is a* ***simple function*** *if ϕ is $\mathcal{M}$ – $\mathcal{B}_{\mathbb{F}}$ measurable and $\phi(X)$ contains only finitely many elements.*

Any such simple functions can be written as

$$\phi = \sum_{i=1}^{n} \lambda_i 1_{A_i} \text{ with } A_i \in \mathcal{M} \text{ and } \lambda_i \in \mathbb{F}. \tag{18.7}$$

Indeed, take $\lambda_1, \lambda_2, \ldots, \lambda_n$ to be an enumeration of the range of ϕ and $A_i = \phi^{-1}(\{\lambda_i\})$. Note that this argument shows that any simple function may be written intrinsically as

$$\phi = \sum_{y \in \mathbb{F}} y 1_{\phi^{-1}(\{y\})}. \tag{18.8}$$

The next theorem shows that simple functions are "pointwise dense" in the space of measurable functions.

Theorem 18.42 (Approximation Theorem). *Let $f : X \to [0,\infty]$ be measurable and define, see Figure 18.2,*

$$\begin{aligned}\phi_n(x) &:= \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{f^{-1}\left(\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]\right)}(x) + 2^n 1_{f^{-1}((2^n,\infty])}(x) \\ &= \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{\left\{\frac{k}{2^n} < f \le \frac{k+1}{2^n}\right\}}(x) + 2^n 1_{\{f > 2^n\}}(x)\end{aligned}$$

then $\phi_n \le f$ for all n, $\phi_n(x) \uparrow f(x)$ for all $x \in X$ and $\phi_n \uparrow f$ uniformly on the sets $X_M := \{x \in X : f(x) \le M\}$ with $M < \infty$. Moreover, if $f : X \to \mathbb{C}$ is a measurable function, then there exists simple functions ϕ_n such that $\lim_{n\to\infty} \phi_n(x) = f(x)$ for all x and $|\phi_n| \uparrow |f|$ as $n \to \infty$.

Proof. Since

$$\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right] = \left(\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}}\right] \cup \left(\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}}\right],$$

if $x \in f^{-1}\left(\left(\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}}\right]\right)$ then $\phi_n(x) = \phi_{n+1}(x) = \frac{2k}{2^{n+1}}$ and if $x \in f^{-1}\left(\left(\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}}\right]\right)$ then $\phi_n(x) = \frac{2k}{2^{n+1}} < \frac{2k+1}{2^{n+1}} = \phi_{n+1}(x)$. Similarly

$$(2^n, \infty] = (2^n, 2^{n+1}] \cup (2^{n+1}, \infty],$$

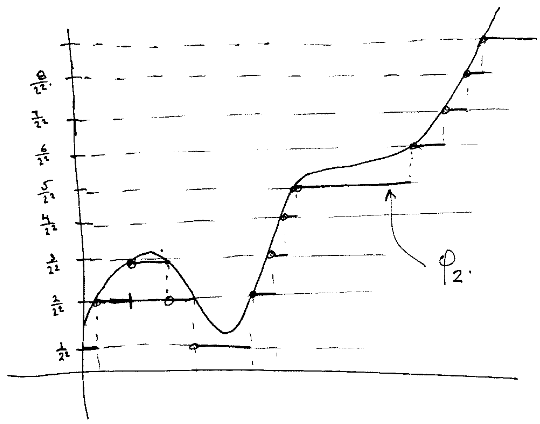


Fig. 18.2. Constructing simple functions approximating a function, $f : X \to [0, \infty]$.

and so for $x \in f^{-1}((2^{n+1}, \infty])$, $\phi_n(x) = 2^n < 2^{n+1} = \phi_{n+1}(x)$ and for $x \in f^{-1}((2^n, 2^{n+1}])$, $\phi_{n+1}(x) \geq 2^n = \phi_n(x)$. Therefore $\phi_n \leq \phi_{n+1}$ for all n. It is clear by construction that $\phi_n(x) \leq f(x)$ for all x and that $0 \leq f(x) - \phi_n(x) \leq 2^{-n}$ if $x \in X_{2^n}$. Hence we have shown that $\phi_n(x) \uparrow f(x)$ for all $x \in X$ and $\phi_n \uparrow f$ uniformly on bounded sets. For the second assertion, first assume that $f : X \to \mathbb{R}$ is a measurable function and choose $\phi_n^{\pm}$ to be simple functions such that $\phi_n^{\pm} \uparrow f_{\pm}$ as $n \to \infty$ and define $\phi_n = \phi_n^+ - \phi_n^-$. Then

$$|\phi_n| = \phi_n^+ + \phi_n^- \leq \phi_{n+1}^+ + \phi_{n+1}^- = |\phi_{n+1}|$$

and clearly $|\phi_n| = \phi_n^+ + \phi_n^- \uparrow f_+ + f_- = |f|$ and $\phi_n = \phi_n^+ - \phi_n^- \to f_+ - f_- = f$ as $n \to \infty$. Now suppose that $f : X \to \mathbb{C}$ is measurable. We may now choose simple function u_n and v_n such that $|u_n| \uparrow |\mathrm{Re}\, f|$, $|v_n| \uparrow |\mathrm{Im}\, f|$, $u_n \to \mathrm{Re}\, f$ and $v_n \to \mathrm{Im}\, f$ as $n \to \infty$. Let $\phi_n = u_n + iv_n$, then

$$|\phi_n|^2 = u_n^2 + v_n^2 \uparrow |\mathrm{Re}\, f|^2 + |\mathrm{Im}\, f|^2 = |f|^2$$

and $\phi_n = u_n + iv_n \to \mathrm{Re}\, f + i\, \mathrm{Im}\, f = f$ as $n \to \infty$. ■

For the rest of this section let X be a given set.

Definition 18.43 (Bounded Convergence). *We say that a sequence of functions f_n from X to $\mathbb{R}$ or $\mathbb{C}$ **converges boundedly** to a function f if $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X$ and*

$$\sup\{|f_n(x)| : x \in X \text{ and } n = 1, 2, \ldots\} < \infty.$$

Definition 18.44. *A **function algebra** $\mathcal{H}$ on X is a linear subspace of $\ell^\infty(X, \mathbb{R})$ which contains 1 and is closed under pointwise multiplication, i.e. $\mathcal{H}$ is a subalgebra of $\ell^\infty(X, \mathbb{R})$ which contains 1. If $\mathcal{H}$ is further closed under bounded convergence then $\mathcal{H}$ is said to be a σ – **function algebra.***

Example 18.45. Suppose $\mathcal{M}$ is a σ – algebra on X, then

$$\ell^{\infty}(\mathcal{M},\mathbb{R}) := \{f \in \ell^{\infty}(X,\mathbb{R}) : \ f \text{ is } \mathcal{M}/\mathcal{B}_{\mathbb{R}} - \text{measurable}\} \tag{18.9}$$

is a σ – function algebra. The next theorem will show that these are the only example of σ – function algebras. (See Exercise 18.7 below for examples of function algebras on X.)

Notation 18.46 *If $\mathcal{H} \subset \ell^{\infty}(X,\mathbb{R})$ be a function algebra, let*

$$\mathcal{M}(\mathcal{H}) := \{A \subset X : 1_A \in \mathcal{H}\}. \tag{18.10}$$

Theorem 18.47. *Let $\mathcal{H}$ be a σ – function algebra on a set X. Then*

1. *$\mathcal{M}(\mathcal{H})$ is a σ – algebra on X.*
2. *$\mathcal{H} = \ell^{\infty}(\mathcal{M}(\mathcal{H}),\mathbb{R})$.*
3. *The map*

$$\mathcal{M} \in \{\sigma - \text{algebras on } X\} \to \ell^{\infty}(\mathcal{M},\mathbb{R}) \in \{\sigma - \text{function algebras on } X\} \tag{18.11}$$

is bijective with inverse given by $\mathcal{H} \to \mathcal{M}(\mathcal{H})$.

Proof. Let $\mathcal{M} := \mathcal{M}(\mathcal{H})$.

1. Since $0,1 \in \mathcal{H}$, $\emptyset, X \in \mathcal{M}$. If $A \in \mathcal{M}$ then, since $\mathcal{H}$ is a linear subspace of $\ell^{\infty}(X,\mathbb{R})$, $1_{A^c} = 1 - 1_A \in \mathcal{H}$ which shows $A^c \in \mathcal{M}$. If $\{A_n\}_{n=1}^{\infty} \subset \mathcal{M}$, then since $\mathcal{H}$ is an algebra,
$$1_{\cap_{n=1}^{N} A_n} = \prod_{n=1}^{N} 1_{A_n} =: f_N \in \mathcal{H}$$
for all $N \in \mathbb{N}$. Because $\mathcal{H}$ is closed under bounded convergence it follows that
$$1_{\cap_{n=1}^{\infty} A_n} = \lim_{N\to\infty} f_N \in \mathcal{H}$$
and this implies $\cap_{n=1}^{\infty} A_n \in \mathcal{M}$. Hence we have shown $\mathcal{M}$ is a σ – algebra.
2. Since $\mathcal{H}$ is an algebra, $p(f) \in \mathcal{H}$ for any $f \in \mathcal{H}$ and any polynomial p on $\mathbb{R}$. The Weierstrass approximation Theorem 8.34, asserts that polynomials on $\mathbb{R}$ are uniformly dense in the space of continuos functions on any compact subinterval of $\mathbb{R}$. Hence if $f \in \mathcal{H}$ and $\phi \in C(\mathbb{R})$, there exists polynomials p_n on $\mathbb{R}$ such that $p_n \circ f(x)$ converges to $\phi \circ f(x)$ uniformly (and hence boundedly) in $x \in X$ as $n \to \infty$. Therefore $\phi \circ f \in \mathcal{H}$ for all $f \in \mathcal{H}$ and $\phi \in C(\mathbb{R})$ and in particular $|f| \in \mathcal{H}$ and $f_{\pm} := \frac{|f| \pm f}{2} \in \mathcal{H}$ if $f \in \mathcal{H}$. Fix an $\alpha \in \mathbb{R}$ and for $n \in \mathbb{N}$ let $\phi_n(t) := (t-\alpha)_+^{1/n}$, where $(t-\alpha)_+ := \max\{t-\alpha, 0\}$. Then $\phi_n \in C(\mathbb{R})$ and $\phi_n(t) \to 1_{t>\alpha}$ as $n \to \infty$ and the convergence is bounded when t is restricted to any compact subset of $\mathbb{R}$. Hence if $f \in \mathcal{H}$ it follows that $1_{f>\alpha} = \lim_{n\to\infty} \phi_n(f) \in \mathcal{H}$ for all $\alpha \in \mathbb{R}$,

i.e. $\{f > \alpha\} \in \mathcal{M}$ for all $\alpha \in \mathbb{R}$. Therefore if $f \in \mathcal{H}$ then $f \in \ell^\infty(\mathcal{M}, \mathbb{R})$ and we have shown $\mathcal{H} \subset \ell^\infty(\mathcal{M}, \mathbb{R})$. Conversely if $f \in \ell^\infty(\mathcal{M}, \mathbb{R})$, then for any $\alpha < \beta$, $\{\alpha < f \le \beta\} \in \mathcal{M} = \mathcal{M}(\mathcal{H})$ and so by assumption $1_{\{\alpha < f \le \beta\}} \in \mathcal{H}$. Combining this remark with the approximation Theorem 18.42 and the fact that $\mathcal{H}$ is closed under bounded convergence shows that $f \in \mathcal{H}$. Hence we have shown $\ell^\infty(\mathcal{M}, \mathbb{R}) \subset \mathcal{H}$ which combined with $\mathcal{H} \subset \ell^\infty(\mathcal{M}, \mathbb{R})$ already proved shows $\mathcal{H} = \ell^\infty(\mathcal{M}(\mathcal{H}), \mathbb{R})$.

3. Items 1. and 2. shows the map in Eq. (18.11) is surjective. To see the map is injective suppose $\mathcal{M}$ and $\mathcal{F}$ are two σ – algebras on X such that $\ell^\infty(\mathcal{M}, \mathbb{R}) = \ell^\infty(\mathcal{F}, \mathbb{R})$, then

$$\begin{aligned}\mathcal{M} &= \{A \subset X : 1_A \in \ell^\infty(\mathcal{M}, \mathbb{R})\} \\ &= \{A \subset X : 1_A \in \ell^\infty(\mathcal{F}, \mathbb{R})\} = \mathcal{F}.\end{aligned}$$

■

Notation 18.48 *Suppose M is a subset of $\ell^\infty(X, \mathbb{R})$.*

1. *Let $\mathcal{H}(M)$ denote the smallest subspace of $\ell^\infty(X, \mathbb{R})$ which contains M and the constant functions and is closed under bounded convergence.*
2. *Let $\mathcal{H}_\sigma(M)$ denote the smallest σ – function algebra containing M.*

Theorem 18.49. *Suppose M is a subset of $\ell^\infty(X, \mathbb{R})$, then $\mathcal{H}_\sigma(M) = \ell^\infty(\sigma(M), \mathbb{R})$ or in other words the following diagram commutes:*

$$\begin{array}{ccccc} & M & \longrightarrow & \sigma(M) & \\ M & \{\text{Multiplicative Subsets}\} & \longrightarrow & \{\sigma - \text{algebras}\} & \mathcal{M} \\ \downarrow & \downarrow & & \downarrow & \downarrow \\ \mathcal{H}_\sigma(M) & \{\sigma\text{– function algebras}\} & = & \{\sigma\text{– function algebras}\} & \ell^\infty(\mathcal{M}, \mathbb{R}). \end{array}$$

Proof. Since $\ell^\infty(\sigma(M), \mathbb{R})$ is σ – function algebra which contains M it follows that

$$\mathcal{H}_\sigma(M) \subset \ell^\infty(\sigma(M), \mathbb{R}).$$

For the opposite inclusion, let

$$\mathcal{M} = \mathcal{M}(\mathcal{H}_\sigma(M)) := \{A \subset X : 1_A \in \mathcal{H}_\sigma(M)\}.$$

By Theorem 18.47, $M \subset \mathcal{H}_\sigma(M) = \ell^\infty(\mathcal{M}, \mathbb{R})$ which implies that every $f \in M$ is $\mathcal{M}$ – measurable. This then implies $\sigma(M) \subset \mathcal{M}$ and therefore

$$\ell^\infty(\sigma(M), \mathbb{R}) \subset \ell^\infty(\mathcal{M}, \mathbb{R}) = \mathcal{H}_\sigma(M).$$

■

Definition 18.50 (Multiplicative System). *A collection of bounded real or complex valued functions, M, on a set X is called a* ***multiplicative system*** *if $f \cdot g \in M$ whenever f and g are in M.*

Theorem 18.51 (Dynkin's Multiplicative System Theorem). *Suppose $M \subset \ell^{\infty}(X, \mathbb{R})$ is a multiplicative system, then*

$$\mathcal{H}(M) = \mathcal{H}_{\sigma}(M) = \ell^{\infty}(\sigma(M), \mathbb{R}). \tag{18.12}$$

In words, the smallest subspace of bounded real valued functions on X which contains M that is closed under bounded convergence is the same as the space of bounded real valued $\sigma(M)$ – measurable functions on X.

Proof. We begin by proving $\mathcal{H} := \mathcal{H}(M)$ is a σ – function algebra. To do this, for any $f \in \mathcal{H}$ let

$$\mathcal{H}_f := \{g \in \mathcal{H} : fg \in \mathcal{H}\} \subset \mathcal{H}$$

and notice that $\mathcal{H}_f$ is a linear subspace of $\ell^{\infty}(X, \mathbb{R})$ which is closed under bounded convergence. Moreover if $f \in M$, $M \subset \mathcal{H}_f$ since M is multiplicative. Therefore $\mathcal{H}_f = \mathcal{H}$ and we have shown that $fg \in \mathcal{H}$ whenever $f \in M$ and $g \in \mathcal{H}$. Given this it now follows that $M \subset \mathcal{H}_f$ for any $f \in \mathcal{H}$ and by the same reasoning just used, $\mathcal{H}_f = \mathcal{H}$. Since $f \in \mathcal{H}$ is arbitrary, we have shown $fg \in \mathcal{H}$ for all $f, g \in \mathcal{H}$, i.e. $\mathcal{H}$ is an algebra. Since it is harder to be an algebra of functions containing M (see Exercise 18.13) than it is to be a subspace of functions containing M it follows that $\mathcal{H}(M) \subset \mathcal{H}_{\sigma}(M)$. But as we have just seen $\mathcal{H}(M)$ is a σ – function algebra which contains M so we must have $\mathcal{H}_{\sigma}(M) \subset \mathcal{H}(M)$ because $\mathcal{H}_{\sigma}(M)$ is by definition the smallest such σ – function algebra. Hence $\mathcal{H}_{\sigma}(M) = \mathcal{H}(M)$. The assertion that $\mathcal{H}_{\sigma}(M) = \ell^{\infty}(\sigma(M), \mathbb{R})$ has already been proved in Theorem 18.49. ■

Theorem 18.52 (Complex Multiplicative System Theorem). *Suppose $\mathcal{H}$ is a complex linear subspace of $\ell^{\infty}(X, \mathbb{C})$ such that: $1 \in \mathcal{H}$, $\mathcal{H}$ is closed under complex conjugation, and $\mathcal{H}$ is closed under bounded convergence. If $M \subset \mathcal{H}$ is multiplicative system which is closed under conjugation, then $\mathcal{H}$ contains all bounded complex valued $\sigma(M)$-measurable functions, i.e. $\ell^{\infty}(\sigma(M), \mathbb{C}) \subset \mathcal{H}$.*

Proof. Let $M_0 = \operatorname{span}_{\mathbb{C}}(M \cup \{1\})$ be the complex span of M. As the reader should verify, M_0 is an algebra, $M_0 \subset \mathcal{H}$, M_0 is closed under complex conjugation and that $\sigma(M_0) = \sigma(M)$. Let $\mathcal{H}^{\mathbb{R}} := \mathcal{H} \cap \ell^{\infty}(X, \mathbb{R})$ and $M_0^{\mathbb{R}} = M \cap \ell^{\infty}(X, \mathbb{R})$. Then (you verify) $M_0^{\mathbb{R}}$ is a multiplicative system, $M_0^{\mathbb{R}} \subset \mathcal{H}^{\mathbb{R}}$ and $\mathcal{H}^{\mathbb{R}}$ is a linear space containing 1 which is closed under bounded convergence. Therefore by Theorem 18.51, $\ell^{\infty}\left(\sigma\left(M_0^{\mathbb{R}}\right), \mathbb{R}\right) \subset \mathcal{H}^{\mathbb{R}}$. Since $\mathcal{H}$ and M_0 are complex linear spaces closed under complex conjugation, for any $f \in \mathcal{H}$ or $f \in M_0$, the functions $\operatorname{Re} f = \frac{1}{2}\left(f + \bar{f}\right)$ and $\operatorname{Im} f = \frac{1}{2i}\left(f - \bar{f}\right)$ are in $\mathcal{H}$ (M_0) or M_0 respectively. Therefore $\mathcal{H} = \mathcal{H}^{\mathbb{R}} + i\mathcal{H}^{\mathbb{R}}$, $M_0 = M_0^{\mathbb{R}} + iM_0^{\mathbb{R}}$, $\sigma\left(M_0^{\mathbb{R}}\right) = \sigma(M_0) = \sigma(M)$ and

$$\begin{aligned}\ell^{\infty}(\sigma(M), \mathbb{C}) &= \ell^{\infty}\left(\sigma\left(M_0^{\mathbb{R}}\right), \mathbb{R}\right) + i\ell^{\infty}\left(\sigma\left(M_0^{\mathbb{R}}\right), \mathbb{R}\right) \\ &\subset \mathcal{H}^{\mathbb{R}} + i\mathcal{H}^{\mathbb{R}} = \mathcal{H}.\end{aligned}$$

■

Exercise 18.7 (Algebra analogue of Theorem 18.47). Call a function algebra $\mathcal{H} \subset \ell^\infty(X, \mathbb{R})$ a **simple function algebra** if the range of each function $f \in \mathcal{H}$ is a finite subset of $\mathbb{R}$. Prove there is a one to one correspondence between algebras $\mathcal{A}$ on a set X and simple function algebras $\mathcal{H}$ on X.

Definition 18.53. *A collection of subsets, $\mathcal{C}$, of X is a* ***multiplicative class****(or a π –* ***class****) if $\mathcal{C}$ is closed under finite intersections.*

Corollary 18.54. *Suppose $\mathcal{H}$ is a subspace of $\ell^\infty(X, \mathbb{R})$ which is closed under bounded convergence and $1 \in \mathcal{H}$. If $\mathcal{C} \subset 2^X$ is a multiplicative class such that $1_A \in \mathcal{H}$ for all $A \in \mathcal{C}$, then $\mathcal{H}$ contains all bounded $\sigma(\mathcal{C})$ – measurable functions.*

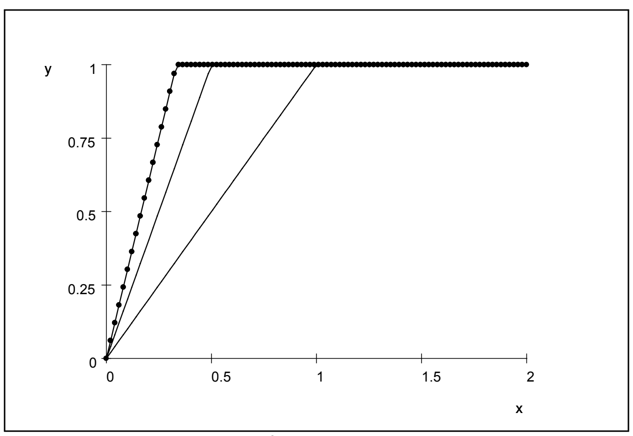
Proof. Let $M = \{1\} \cup \{1_A : A \in \mathcal{C}\}$. Then $M \subset \mathcal{H}$ is a multiplicative system and the proof is completed with an application of Theorem 18.51. ■

Corollary 18.55. *Suppose that (X, d) is a metric space and $\mathcal{B}_X = \sigma(\tau_d)$ is the Borel σ – algebra on X and $\mathcal{H}$ is a subspace of $\ell^\infty(X, \mathbb{R})$ such that $BC(X, \mathbb{R}) \subset \mathcal{H}$ and $\mathcal{H}$ is closed under bounded convergence*[1]*. Then $\mathcal{H}$ contains all bounded $\mathcal{B}_X$ – measurable real valued functions on X. (This may be stated as follows: the smallest vector space of bounded functions which is closed under bounded convergence and contains $BC(X, \mathbb{R})$ is the space of bounded $\mathcal{B}_X$ – measurable real valued functions on X.)*

Proof. Let $V \in \tau_d$ be an open subset of X and for $n \in \mathbb{N}$ let

$$f_n(x) := \min(n \cdot d_{V^c}(x), 1) \text{ for all } x \in X.$$

Notice that $f_n = \phi_n \circ d_{V^c}$ where $\phi_n(t) = \min(nt, 1)$ (see Figure 18.3) which is continuous and hence $f_n \in BC(X, \mathbb{R})$ for all n. Furthermore, f_n converges boundedly to $1_{d_{V^c} > 0} = 1_V$ as $n \to \infty$ and therefore $1_V \in \mathcal{H}$ for all $V \in \tau$. Since τ is a π – class, the result now follows by an application of Corollary 18.54.



Plots of ϕ_1, ϕ_2 and ϕ_3.

[1] Recall that $BC(X, \mathbb{R})$ are the bounded continuous functions on X.

■

Here are some more variants of Corollary 18.55.

Proposition 18.56. *Let (X, d) be a metric space, $\mathcal{B}_X = \sigma(\tau_d)$ be the Borel σ – algebra and assume there exists compact sets $K_k \subset X$ such that $K_k^o \uparrow X$. Suppose that $\mathcal{H}$ is a subspace of $\ell^\infty(X, \mathbb{R})$ such that $C_c(X, \mathbb{R}) \subset \mathcal{H}$ ($C_c(X, \mathbb{R})$ is the space of continuous functions with compact support) and $\mathcal{H}$ is closed under bounded convergence. Then $\mathcal{H}$ contains all bounded $\mathcal{B}_X$ – measurable real valued functions on X.*

Proof. Let k and n be positive integers and set $\psi_{n,k}(x) = \min(1, n \cdot d_{(K_k^o)^c}(x))$. Then $\psi_{n,k} \in C_c(X, \mathbb{R})$ and $\{\psi_{n,k} \neq 0\} \subset K_k^o$. Let $\mathcal{H}_{n,k}$ denote those bounded $\mathcal{B}_X$ – measurable functions, $f : X \to \mathbb{R}$, such that $\psi_{n,k} f \in \mathcal{H}$. It is easily seen that $\mathcal{H}_{n,k}$ is closed under bounded convergence and that $\mathcal{H}_{n,k}$ contains $BC(X, \mathbb{R})$ and therefore by Corollary 18.55, $\psi_{n,k} f \in \mathcal{H}$ for all bounded measurable functions $f : X \to \mathbb{R}$. Since $\psi_{n,k} f \to 1_{K_k^o} f$ boundedly as $n \to \infty$, $1_{K_k^o} f \in \mathcal{H}$ for all k and similarly $1_{K_k^o} f \to f$ boundedly as $k \to \infty$ and therefore $f \in \mathcal{H}$. ■

Lemma 18.57. *Suppose that (X, τ) is a locally compact second countable Hausdorff space.*[2] *Then:*

1. *every open subset $U \subset X$ is σ – compact.*
2. *If $F \subset X$ is a closed set, there exist open sets $V_n \subset X$ such that $V_n \downarrow F$ as $n \to \infty$.*
3. *To each open set $U \subset X$ there exists $f_n \prec U$ (i.e. $f_n \in C_c(U, [0, 1])$) such that $\lim_{n\to\infty} f_n = 1_U$.*
4. *$\mathcal{B}_X = \sigma(C_c(X, \mathbb{R}))$, i.e. the σ – algebra generated by $C_c(X)$ is the Borel σ – algebra on X.*

Proof.

1. Let U be an open subset of X, $\mathcal{V}$ be a countable base for τ and
$$\mathcal{V}^U := \{W \in \mathcal{V} : \bar{W} \subset U \text{ and } \bar{W} \text{ is compact}\}.$$
For each $x \in U$, by Proposition 12.7, there exists an open neighborhood V of x such that $\bar{V} \subset U$ and $\bar{V}$ is compact. Since $\mathcal{V}$ is a base for the topology τ, there exists $W \in \mathcal{V}$ such that $x \in W \subset V$. Because $\bar{W} \subset \bar{V}$, it follows that $\bar{W}$ is compact and hence $W \in \mathcal{V}^U$. As $x \in U$ was arbitrary, $U = \cup \mathcal{V}^U$. Let $\{W_n\}_{n=1}^\infty$ be an enumeration of $\mathcal{V}^U$ and set $K_n := \cup_{k=1}^n \bar{W}_k$. Then $K_n \uparrow U$ as $n \to \infty$ and K_n is compact for each n.
2. Let $\{K_n\}_{n=1}^\infty$ be compact subsets of F^c such that $K_n \uparrow F^c$ as $n \to \infty$ and set $V_n := K_n^c = X \setminus K_n$. Then $V_n \downarrow F$ and by Proposition 12.5, V_n is open for each n.

[2] For example any separable locally compact metric space and in particular any open subset of $\mathbb{R}^n$.

3. Let $U \subset X$ be an open set and $\{K_n\}_{n=1}^\infty$ be compact subsets of U such that $K_n \uparrow U$. By Lemma 12.8, there exist $f_n \prec U$ such that $f_n = 1$ on K_n. These functions satisfy, $1_U = \lim_{n\to\infty} f_n$.
4. By item 3., 1_U is $\sigma(C_c(X,\mathbb{R}))$ – measurable for all $U \in \tau$ and hence $\tau \subset \sigma(C_c(X,\mathbb{R}))$. Therefore $\mathcal{B}_X = \sigma(\tau) \subset \sigma(C_c(X,\mathbb{R}))$. The converse inclusion holds because continuous functions are always Borel measurable.

■

Here is a variant of Corollary 18.55.

Corollary 18.58. *Suppose that (X,τ) is a second countable locally compact Hausdorff space and $\mathcal{B}_X = \sigma(\tau)$ is the Borel σ – algebra on X. If $\mathcal{H}$ is a subspace of $\ell^\infty(X,\mathbb{R})$ which is closed under bounded convergence and contains $C_c(X,\mathbb{R})$, then $\mathcal{H}$ contains all bounded $\mathcal{B}_X$ – measurable real valued functions on X.*

Proof. By Item 3. of Lemma 18.57, for every $U \in \tau$ the characteristic function, 1_U, may be written as a bounded pointwise limit of functions from $C_c(X,\mathbb{R})$. Therefore $1_U \in \mathcal{H}$ for all $U \in \tau$. Since τ is a π – class, the proof is finished with an application of Corollary 18.54 ■

18.4 Product σ – Algebras

Let $\{(X_\alpha, \mathcal{M}_\alpha)\}_{\alpha\in A}$ be a collection of measurable spaces $X = X_A = \prod_{\alpha\in A} X_\alpha$ and $\pi_\alpha : X_A \to X_\alpha$ be the canonical projection map as in Notation 2.2.

Definition 18.59 (Product σ – Algebra). *The **product σ – algebra,** $\otimes_{\alpha\in A}\mathcal{M}_\alpha$, is the smallest σ – algebra on X such that each π_α for $\alpha \in A$ is measurable, i.e.*

$$\otimes_{\alpha\in A}\mathcal{M}_\alpha := \sigma(\pi_\alpha : \alpha \in A) = \sigma\left(\cup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)\right).$$

Applying Proposition 18.25 in this setting implies the following proposition.

Proposition 18.60. *Suppose Y is a measurable space and $f : Y \to X = X_A$ is a map. Then f is measurable iff $\pi_\alpha \circ f : Y \to X_\alpha$ is measurable for all $\alpha \in A$. In particular if $A = \{1, 2, \ldots, n\}$ so that $X = X_1 \times X_2 \times \cdots \times X_n$ and $f(y) = (f_1(y), f_2(y), \ldots, f_n(y)) \in X_1 \times X_2 \times \cdots \times X_n$, then $f : Y \to X_A$ is measurable iff $f_i : Y \to X_i$ is measurable for all i.*

Proposition 18.61. *Suppose that $(X_\alpha, \mathcal{M}_\alpha)_{\alpha\in A}$ is a collection of measurable spaces and $\mathcal{E}_\alpha \subset \mathcal{M}_\alpha$ generates $\mathcal{M}_\alpha$ for each $\alpha \in A$, then*

$$\otimes_{\alpha\in A}\mathcal{M}_\alpha = \sigma\left(\cup_{\alpha\in A} \pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \tag{18.13}$$

Moreover, suppose that A is either finite or countably infinite, $X_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$, and $\mathcal{M}_\alpha = \sigma(\mathcal{E}_\alpha)$ for each $\alpha \in A$. Then the product σ – algebra satisfies

$$\otimes_{\alpha\in A}\mathcal{M}_\alpha = \sigma\left(\left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}\right). \tag{18.14}$$

In particular if $A = \{1, 2, \ldots, n\}$, then $X = X_1 \times X_2 \times \cdots \times X_n$ and

$$\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \cdots \otimes \mathcal{M}_n = \sigma(\mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_n),$$

where $\mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_n$ is as defined in Notation 10.26.

Proof. Since $\cup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha) \subset \cup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)$, it follows that

$$\mathcal{F} := \sigma\left(\cup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \subset \sigma\left(\cup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)\right) = \otimes_{\alpha\in A}\mathcal{M}_\alpha.$$

Conversely,

$$\mathcal{F} \supset \sigma(\pi_\alpha^{-1}(\mathcal{E}_\alpha)) = \pi_\alpha^{-1}(\sigma(\mathcal{E}_\alpha)) = \pi_\alpha^{-1}(\mathcal{M}_\alpha)$$

holds for all α implies that

$$\cup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha) \subset \mathcal{F}$$

and hence that $\otimes_{\alpha\in A}\mathcal{M}_\alpha \subset \mathcal{F}$. We now prove Eq. (18.14). Since we are assuming that $X_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$, we see that

$$\cup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha) \subset \left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}$$

and therefore by Eq. (18.13)

$$\otimes_{\alpha\in A}\mathcal{M}_\alpha = \sigma\left(\cup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \subset \sigma\left(\left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}\right).$$

This last statement is true independent as to whether A is countable or not. For the reverse inclusion it suffices to notice that since A is countable,

$$\prod_{\alpha\in A} E_\alpha = \cap_{\alpha\in A}\pi_\alpha^{-1}(E_\alpha) \in \otimes_{\alpha\in A}\mathcal{M}_\alpha$$

and hence

$$\sigma\left(\left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}\right) \subset \otimes_{\alpha\in A}\mathcal{M}_\alpha.$$

■

Remark 18.62. One can not relax the assumption that $X_\alpha \in \mathcal{E}_\alpha$ in the second part of Proposition 18.61. For example, if $X_1 = X_2 = \{1,2\}$ and $\mathcal{E}_1 = \mathcal{E}_2 = \{\{1\}\}$, then $\sigma(\mathcal{E}_1 \times \mathcal{E}_2) = \{\emptyset, X_1 \times X_2, \{(1,1)\}\}$ while $\sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2)) = 2^{X_1 \times X_2}$.

Theorem 18.63. *Let $\{X_\alpha\}_{\alpha \in A}$ be a sequence of sets where A is at most countable. Suppose for each $\alpha \in A$ we are given a countable set $\mathcal{E}_\alpha \subset 2^{X_\alpha}$. Let $\tau_\alpha = \tau(\mathcal{E}_\alpha)$ be the topology on X_α generated by $\mathcal{E}_\alpha$ and X be the product space $\prod_{\alpha \in A} X_\alpha$ with equipped with the product topology $\tau := \otimes_{\alpha \in A} \tau(\mathcal{E}_\alpha)$. Then the Borel σ – algebra $\mathcal{B}_X = \sigma(\tau)$ is the same as the product σ – algebra:*

$$\mathcal{B}_X = \otimes_{\alpha \in A} \mathcal{B}_{X_\alpha},$$

where $\mathcal{B}_{X_\alpha} = \sigma(\tau(\mathcal{E}_\alpha)) = \sigma(\mathcal{E}_\alpha)$ for all $\alpha \in A$.

In particular if $A = \{1, 2, \ldots, n\}$ and each (X_i, τ_i) is a second countable topological space, then

$$\mathcal{B}_X := \sigma(\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n) = \sigma(\mathcal{B}_{X_1} \times \cdots \times \mathcal{B}_{X_n}) =: \mathcal{B}_{X_1} \otimes \cdots \otimes \mathcal{B}_{X_n}.$$

Proof. By Proposition 10.25, the topology τ may be described as the smallest topology containing $\mathcal{E} = \cup_{\alpha \in A} \pi_\alpha^{-1}(\mathcal{E}_\alpha)$. Now $\mathcal{E}$ is the countable union of countable sets so is still countable. Therefore by Proposition 18.17 and Proposition 18.61,

$$\begin{aligned} \mathcal{B}_X &= \sigma(\tau) = \sigma(\tau(\mathcal{E})) = \sigma(\mathcal{E}) = \otimes_{\alpha \in A} \sigma(\mathcal{E}_\alpha) \\ &= \otimes_{\alpha \in A} \sigma(\tau_\alpha) = \otimes_{\alpha \in A} \mathcal{B}_{X_\alpha}. \end{aligned}$$

■

Corollary 18.64. *If (X_i, d_i) are separable metric spaces for $i = 1, \ldots, n$, then*

$$\mathcal{B}_{X_1} \otimes \cdots \otimes \mathcal{B}_{X_n} = \mathcal{B}_{(X_1 \times \cdots \times X_n)}$$

where $\mathcal{B}_{X_i}$ is the Borel σ – algebra on X_i and $\mathcal{B}_{(X_1 \times \cdots \times X_n)}$ is the Borel σ – algebra on $X_1 \times \cdots \times X_n$ equipped with the metric topology associated to the metric $d(x,y) = \sum_{i=1}^n d_i(x_i, y_i)$ where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$.

Proof. This is a combination of the results in Lemma 10.28, Exercise 10.10 and Theorem 18.63. ■

Because all norms on finite dimensional spaces are equivalent, the usual Euclidean norm on $\mathbb{R}^m \times \mathbb{R}^n$ is equivalent to the “product” norm defined by

$$\|(x,y)\|_{\mathbb{R}^m \times \mathbb{R}^n} = \|x\|_{\mathbb{R}^m} + \|y\|_{\mathbb{R}^n}.$$

Hence by Lemma 10.28, the Euclidean topology on $\mathbb{R}^{m+n}$ is the same as the product topology on $\mathbb{R}^{m+n} \cong \mathbb{R}^m \times \mathbb{R}^n$. Here we are identifying $\mathbb{R}^m \times \mathbb{R}^n$ with $\mathbb{R}^{m+n}$ by the map

$$(x,y) \in \mathbb{R}^m \times \mathbb{R}^n \to (x_1, \ldots, x_m, y_1, \ldots, y_n) \in \mathbb{R}^{m+n}.$$

These comments along with Corollary 18.64 proves the following result.

Corollary 18.65. *After identifying $\mathbb{R}^m \times \mathbb{R}^n$ with $\mathbb{R}^{m+n}$ as above and letting $\mathcal{B}_{\mathbb{R}^n}$ denote the Borel σ –algebra on $\mathbb{R}^n$, we have*

$$\mathcal{B}_{\mathbb{R}^{m+n}} = \mathcal{B}_{\mathbb{R}^n} \otimes \mathcal{B}_{\mathbb{R}^m} \text{ and } \mathcal{B}_{\mathbb{R}^n} = \overbrace{\mathcal{B}_{\mathbb{R}} \otimes \cdots \otimes \mathcal{B}_{\mathbb{R}}}^{n\text{–times}}.$$

18.4.1 Factoring of Measurable Maps

Lemma 18.66. *Suppose that $(Y, \mathcal{F})$ is a measurable space and $F : X \to Y$ is a map. Then to every $(\sigma(F), \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function, $H : X \to \bar{\mathbb{R}}$, there is a $(\mathcal{F}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function $h : Y \to \bar{\mathbb{R}}$ such that $H = h \circ F$.*

Proof. First suppose that $H = 1_A$ where $A \in \sigma(F) = F^{-1}(\mathcal{F})$. Let $B \in \mathcal{F}$ such that $A = F^{-1}(B)$ then $1_A = 1_{F^{-1}(B)} = 1_B \circ F$ and hence the Lemma is valid in this case with $h = 1_B$. More generally if $H = \sum a_i 1_{A_i}$ is a simple function, then there exists $B_i \in \mathcal{F}$ such that $1_{A_i} = 1_{B_i} \circ F$ and hence $H = h \circ F$ with $h := \sum a_i 1_{B_i}$ – a simple function on $\bar{\mathbb{R}}$. For general $(\sigma(F), \mathcal{F})$ – measurable function, H, from $X \to \bar{\mathbb{R}}$, choose simple functions H_n converging to H. Let h_n be simple functions on $\bar{\mathbb{R}}$ such that $H_n = h_n \circ F$. Then it follows that

$$H = \lim_{n\to\infty} H_n = \limsup_{n\to\infty} H_n = \limsup_{n\to\infty} h_n \circ F = h \circ F$$

where $h := \limsup_{n\to\infty} h_n$ – a measurable function from Y to $\bar{\mathbb{R}}$. ■

The following is an immediate corollary of Proposition 18.25 and Lemma 18.66.

Corollary 18.67. *Let X and A be sets, and suppose for $\alpha \in A$ we are give a measurable space $(Y_\alpha, \mathcal{F}_\alpha)$ and a function $f_\alpha : X \to Y_\alpha$. Let $Y := \prod_{\alpha\in A} Y_\alpha$, $\mathcal{F} := \otimes_{\alpha\in A}\mathcal{F}_\alpha$ be the product σ – algebra on Y and $\mathcal{M} := \sigma(f_\alpha : \alpha \in A)$ be the smallest σ – algebra on X such that each f_α is measurable. Then the function $F : X \to Y$ defined by $[F(x)]_\alpha := f_\alpha(x)$ for each $\alpha \in A$ is $(\mathcal{M}, \mathcal{F})$ – measurable and a function $H : X \to \bar{\mathbb{R}}$ is $(\mathcal{M}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable iff there exists a $(\mathcal{F}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function h from Y to $\bar{\mathbb{R}}$ such that $H = h \circ F$.*

18.5 Exercises

Exercise 18.8. Prove Corollary 18.23. **Hint:** See Exercise 18.3.

Exercise 18.9. If $\mathcal{M}$ is the σ – algebra generated by $\mathcal{E} \subset 2^X$, then $\mathcal{M}$ is the union of the σ – algebras generated by countable subsets $\mathcal{F} \subset \mathcal{E}$. (Folland, Problem 1.5 on p.24.)

Exercise 18.10. Let $(X, \mathcal{M})$ be a measure space and $f_n : X \to \mathbb{F}$ be a sequence of measurable functions on X. Show that $\{x : \lim_{n\to\infty} f_n(x) \text{ exists } \mathbf{in}\ \mathbb{F}\} \in \mathcal{M}$.

Exercise 18.11. Show that every monotone function $f : \mathbb{R} \to \mathbb{R}$ is $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}})$ – measurable.

Exercise 18.12. Show by example that the supremum of an uncountable family of measurable functions need not be measurable. (Folland problem 2.6 on p. 48.)

Exercise 18.13. Let $X = \{1, 2, 3, 4\}$, $A = \{1, 2\}$, $B = \{2, 3\}$ and $M := \{1_A, 1_B\}$. Show $\mathcal{H}_\sigma(M) \neq \mathcal{H}(M)$ in this case.

19

Measures and Integration

Definition 19.1. *A* ***measure*** μ *on a measurable space* $(X, \mathcal{M})$ *is a function* $\mu : \mathcal{M} \to [0, \infty]$ *such that*

1. $\mu(\emptyset) = 0$ *and*
2. *(Finite Additivity) If* $\{A_i\}_{i=1}^n \subset \mathcal{M}$ *are pairwise disjoint, i.e.* $A_i \cap A_j = \emptyset$ *when* $i \neq j$*, then*
$$\mu(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i).$$
3. *(Continuity) If* $A_n \in \mathcal{M}$ *and* $A_n \uparrow A$*, then* $\mu(A_n) \uparrow \mu(A)$*.*

We call a triple $(X, \mathcal{M}, \mu)$*, where* $(X, \mathcal{M})$ *is a measurable space and* $\mu : \mathcal{M} \to [0, \infty]$ *is a measure, a* ***measure space.***

Remark 19.2. Properties 2) and 3) in Definition 19.1 are equivalent to the following condition. If $\{A_i\}_{i=1}^\infty \subset \mathcal{M}$ are pairwise disjoint then

$$\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i). \tag{19.1}$$

To prove this assume that Properties 2) and 3) in Definition 19.1 hold and $\{A_i\}_{i=1}^\infty \subset \mathcal{M}$ are pairwise disjoint. Letting $B_n := \bigcup_{i=1}^n A_i \uparrow B := \bigcup_{i=1}^\infty A_i$, we have

$$\mu(\bigcup_{i=1}^\infty A_i) = \mu(B) \overset{(3)}{=} \lim_{n\to\infty} \mu(B_n) \overset{(2)}{=} \lim_{n\to\infty} \sum_{i=1}^n \mu(A_i) = \sum_{i=1}^\infty \mu(A_i).$$

Conversely, if Eq. (19.1) holds we may take $A_j = \emptyset$ for all $j > n$ to see that Property 2) of Definition 19.1 holds. Also if $A_n \uparrow A$, let $B_n := A_n \setminus A_{n-1}$ with $A_0 := \emptyset$. Then $\{B_n\}_{n=1}^\infty$ are pairwise disjoint, $A_n = \cup_{j=1}^n B_j$ and $A = \cup_{j=1}^\infty B_j$. So if Eq. (19.1) holds we have

$$\mu(A) = \mu\left(\cup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} \mu(B_j)$$
$$= \lim_{n\to\infty} \sum_{j=1}^{n} \mu(B_j) = \lim_{n\to\infty} \mu(\cup_{j=1}^{n} B_j) = \lim_{n\to\infty} \mu(A_n).$$

Proposition 19.3 (Basic properties of measures). *Suppose that $(X, \mathcal{M}, \mu)$ is a measure space and $E, F \in \mathcal{M}$ and $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$, then :*

1. *$\mu(E) \le \mu(F)$ if $E \subset F$.*
2. *$\mu(\cup E_j) \le \sum \mu(E_j)$.*
3. *If $\mu(E_1) < \infty$ and $E_j \downarrow E$, i.e. $E_1 \supset E_2 \supset E_3 \supset \dots$ and $E = \cap_j E_j$, then $\mu(E_j) \downarrow \mu(E)$ as $j \to \infty$.*

Proof.

1. Since $F = E \cup (F \setminus E)$,
$$\mu(F) = \mu(E) + \mu(F \setminus E) \ge \mu(E).$$
2. Let $\widetilde{E}_j = E_j \setminus (E_1 \cup \dots \cup E_{j-1})$ so that the $\tilde{E}_j$'s are pair-wise disjoint and $E = \cup \widetilde{E}_j$. Since $\tilde{E}_j \subset E_j$ it follows from Remark 19.2 and part (1), that
$$\mu(E) = \sum \mu(\widetilde{E}_j) \le \sum \mu(E_j).$$
3. Define $D_i := E_1 \setminus E_i$ then $D_i \uparrow E_1 \setminus E$ which implies that
$$\mu(E_1) - \mu(E) = \lim_{i\to\infty} \mu(D_i) = \mu(E_1) - \lim_{i\to\infty} \mu(E_i)$$
which shows that $\lim_{i\to\infty} \mu(E_i) = \mu(E)$.

■

Definition 19.4. *A set $E \subset X$ is a **null** set if $E \in \mathcal{M}$ and $\mu(E) = 0$. If P is some "property" which is either true or false for each $x \in X$, we will use the terminology P a.e. (to be read P almost everywhere) to mean*

$$E := \{x \in X : P \text{ is false for } x\}$$

is a null set. For example if f and g are two measurable functions on $(X, \mathcal{M}, \mu)$, $f = g$ a.e. means that $\mu(f \ne g) = 0$.

Definition 19.5. *A measure space $(X, \mathcal{M}, \mu)$ is **complete** if every subset of a null set is in $\mathcal{M}$, i.e. for all $F \subset X$ such that $F \subset E \in \mathcal{M}$ with $\mu(E) = 0$ implies that $F \in \mathcal{M}$.*

Proposition 19.6. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Set*

$$\mathcal{N} := \{N \subset X : \exists\, F \in \mathcal{M} \ni N \subset F \text{ and } \mu(F) = 0\}$$

and

$$\bar{\mathcal{M}} = \{A \cup N : A \in \mathcal{M}, N \in \mathcal{N}\},$$

see Fig. 19.1. Then $\bar{\mathcal{M}}$ is a σ – algebra. Define $\bar{\mu}(A \cup N) = \mu(A)$, then $\bar{\mu}$ is the unique measure on $\bar{\mathcal{M}}$ which extends μ.

Proof. Clearly $X, \emptyset \in \bar{\mathcal{M}}$. Let $A \in \mathcal{M}$ and $N \in \mathcal{N}$ and choose $F \in \mathcal{M}$

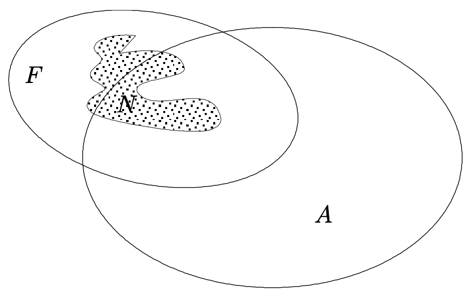


Fig. 19.1. Completing a σ – algebra.

such that $N \subset F$ and $\mu(F) = 0$. Since $N^c = (F \setminus N) \cup F^c$,

$$\begin{aligned}(A \cup N)^c &= A^c \cap N^c = A^c \cap (F \setminus N \cup F^c) \\ &= [A^c \cap (F \setminus N)] \cup [A^c \cap F^c]\end{aligned}$$

where $[A^c \cap (F \setminus N)] \in \mathcal{N}$ and $[A^c \cap F^c] \in \mathcal{M}$. Thus $\bar{\mathcal{M}}$ is closed under complements. If $A_i \in \mathcal{M}$ and $N_i \subset F_i \in \mathcal{M}$ such that $\mu(F_i) = 0$ then $\cup(A_i \cup N_i) = (\cup A_i) \cup (\cup N_i) \in \bar{\mathcal{M}}$ since $\cup A_i \in \mathcal{M}$ and $\cup N_i \subset \cup F_i$ and $\mu(\cup F_i) \leq \sum \mu(F_i) = 0$. Therefore, $\bar{\mathcal{M}}$ is a σ – algebra. Suppose $A \cup N_1 = B \cup N_2$ with $A, B \in \mathcal{M}$ and $N_1, N_2, \in \mathcal{N}$. Then $A \subset A \cup N_1 \subset A \cup N_1 \cup F_2 = B \cup F_2$ which shows that

$$\mu(A) \leq \mu(B) + \mu(F_2) = \mu(B).$$

Similarly, we show that $\mu(B) \leq \mu(A)$ so that $\mu(A) = \mu(B)$ and hence $\bar{\mu}(A \cup N) := \mu(A)$ is well defined. It is left as an exercise to show $\bar{\mu}$ is a measure, i.e. that it is countable additive. ■

Many theorems in the sequel will require some control on the size of a measure μ. The relevant notion for our purposes (and most purposes) is that of a σ – finite measure defined next.

Definition 19.7. *Suppose X is a set, $\mathcal{E} \subset \mathcal{M} \subset 2^X$ and $\mu : \mathcal{M} \to [0, \infty]$ is a function. The function μ is σ – finite on $\mathcal{E}$ if there exists $E_n \in \mathcal{E}$ such that*

$\mu(E_n) < \infty$ and $X = \cup_{n=1}^{\infty} E_n$. If $\mathcal{M}$ is a σ – algebra and μ is a measure on $\mathcal{M}$ which is σ – finite on $\mathcal{M}$ we will say $(X, \mathcal{M}, \mu)$ is a σ – finite measure space.

The reader should check that if μ is a finitely additive measure on an algebra, $\mathcal{M}$, then μ is σ – finite on $\mathcal{M}$ iff there exists $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$.

19.1 Example of Measures

Most σ – algebras and σ -additive measures are somewhat difficult to describe and define. However, one special case is fairly easy to understand. Namely suppose that $\mathcal{F} \subset 2^X$ is a countable or finite partition of X and $\mathcal{M} \subset 2^X$ is the σ – algebra which consists of the collection of sets $A \subset X$ such that

$$A = \cup \{\alpha \in \mathcal{F} : \alpha \subset A\}. \tag{19.2}$$

It is easily seen that $\mathcal{M}$ is a σ – algebra.

Any measure $\mu : \mathcal{M} \to [0, \infty]$ is determined uniquely by its values on $\mathcal{F}$. Conversely, if we are given any function $\lambda : \mathcal{F} \to [0, \infty]$ we may define, for $A \in \mathcal{M}$,

$$\mu(A) = \sum_{\alpha \in \mathcal{F} \ni \alpha \subset A} \lambda(\alpha) = \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A}$$

where $1_{\alpha \subset A}$ is one if $\alpha \subset A$ and zero otherwise. We may check that μ is a measure on $\mathcal{M}$. Indeed, if $A = \coprod_{i=1}^{\infty} A_i$ and $\alpha \in \mathcal{F}$, then $\alpha \subset A$ iff $\alpha \subset A_i$ for one and hence exactly one A_i. Therefore $1_{\alpha \subset A} = \sum_{i=1}^{\infty} 1_{\alpha \subset A_i}$ and hence

$$\begin{aligned} \mu(A) &= \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A} = \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) \sum_{i=1}^{\infty} 1_{\alpha \subset A_i} \\ &= \sum_{i=1}^{\infty} \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A_i} = \sum_{i=1}^{\infty} \mu(A_i) \end{aligned}$$

as desired. Thus we have shown that there is a one to one correspondence between measures μ on $\mathcal{M}$ and functions $\lambda : \mathcal{F} \to [0, \infty]$.

The construction of measures will be covered in Chapters 27 – 29 below. However, let us record here the existence of an interesting class of measures.

Theorem 19.8. *To every right continuous non-decreasing function $F : \mathbb{R} \to \mathbb{R}$ there exists a unique measure μ_F on $\mathcal{B}_{\mathbb{R}}$ such that*

$$\mu_F((a, b]) = F(b) - F(a) \ \forall \ -\infty < a \le b < \infty \tag{19.3}$$

Moreover, if $A \in \mathcal{B}_{\mathbb{R}}$ then

$$\mu_F(A) = \inf\left\{\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) : A \subset \cup_{i=1}^{\infty}(a_i, b_i]\right\} \tag{19.4}$$

$$= \inf\left\{\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) : A \subset \coprod_{i=1}^{\infty}(a_i, b_i]\right\}. \tag{19.5}$$

In fact the map $F \to \mu_F$ is a one to one correspondence between right continuous functions F with $F(0) = 0$ on one hand and measures μ on $\mathcal{B}_{\mathbb{R}}$ such that $\mu(J) < \infty$ on any bounded set $J \in \mathcal{B}_{\mathbb{R}}$ on the other.

Proof. This follows directly from Proposition 26.18 and Theorem 26.2 below. It can also be easily derived from Theorem 26.26 below. ■

Example 19.9. The most important special case of Theorem 19.8 is when $F(x) = x$, in which case we write m for μ_F. The measure m is called Lebesgue measure.

Theorem 19.10. *Lebesgue measure m is invariant under translations, i.e. for $B \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}$,*

$$m(x + B) = m(B). \tag{19.6}$$

Moreover, m is the unique measure on $\mathcal{B}_{\mathbb{R}}$ such that $m((0,1]) = 1$ and Eq. (19.6) holds for $B \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}$. Moreover, m has the scaling property

$$m(\lambda B) = |\lambda|\, m(B) \tag{19.7}$$

where $\lambda \in \mathbb{R}$, $B \in \mathcal{B}_{\mathbb{R}}$ and $\lambda B := \{\lambda x : x \in B\}$.

Proof. Let $m_x(B) := m(x + B)$, then one easily shows that m_x is a measure on $\mathcal{B}_{\mathbb{R}}$ such that $m_x((a,b]) = b - a$ for all $a < b$. Therefore, $m_x = m$ by the uniqueness assertion in Theorem 19.8. For the converse, suppose that m is translation invariant and $m((0,1]) = 1$. Given $n \in \mathbb{N}$, we have

$$(0,1] = \cup_{k=1}^{n}(\frac{k-1}{n}, \frac{k}{n}] = \cup_{k=1}^{n}\left(\frac{k-1}{n} + (0, \frac{1}{n}]\right).$$

Therefore,

$$\begin{aligned} 1 = m((0,1]) &= \sum_{k=1}^{n} m\left(\frac{k-1}{n} + (0, \frac{1}{n}]\right) \\ &= \sum_{k=1}^{n} m((0, \frac{1}{n}]) = n \cdot m((0, \frac{1}{n}]). \end{aligned}$$

That is to say

$$m((0, \frac{1}{n}]) = 1/n.$$

Similarly, $m((0, \frac{l}{n}]) = l/n$ for all $l, n \in \mathbb{N}$ and therefore by the translation invariance of m,

$$m((a,b]) = b - a \text{ for all } a, b \in \mathbb{Q} \text{ with } a < b.$$

Finally for $a, b \in \mathbb{R}$ such that $a < b$, choose $a_n, b_n \in \mathbb{Q}$ such that $b_n \downarrow b$ and $a_n \uparrow a$, then $(a_n, b_n] \downarrow (a, b]$ and thus

$$m((a,b]) = \lim_{n\to\infty} m((a_n, b_n]) = \lim_{n\to\infty} (b_n - a_n) = b - a,$$

i.e. m is Lebesgue measure. To prove Eq. (19.7) we may assume that $\lambda \neq 0$ since this case is trivial to prove. Now let $m_\lambda(B) := |\lambda|^{-1} m(\lambda B)$. It is easily checked that m_λ is again a measure on $\mathcal{B}_{\mathbb{R}}$ which satisfies

$$m_\lambda((a,b]) = \lambda^{-1} m\left((\lambda a, \lambda b]\right) = \lambda^{-1}(\lambda b - \lambda a) = b - a$$

if $\lambda > 0$ and

$$m_\lambda((a,b]) = |\lambda|^{-1} m\left([\lambda b, \lambda a)\right) = -|\lambda|^{-1}(\lambda b - \lambda a) = b - a$$

if $\lambda < 0$. Hence $m_\lambda = m$. ■

We are now going to develop integration theory relative to a measure. The integral defined in the case for Lebesgue measure, m, will be an extension of the standard Riemann integral on $\mathbb{R}$.

19.1.1 ADD: Examples of Measures

BRUCE: ADD details.

1. Product measure for the flipping of a coin.
2. Haar Measure
3. Measure on embedded submanifolds, i.e. Hausdorff measure.
4. Wiener measure.
5. Gibbs states.
6. Measure associated to self-adjoint operators and classifying them.

19.2 Integrals of Simple functions

Let $(X, \mathcal{M}, \mu)$ be a fixed measure space in this section.

Definition 19.11. *Let $\mathbb{F} = \mathbb{C}$ or $[0,\infty)$ and suppose that $\phi : X \to \mathbb{F}$ is a simple function as in Definition 18.41. If $\mathbb{F} = \mathbb{C}$ assume further that $\mu(\phi^{-1}(\{y\})) < \infty$ for all $y \neq 0$ in $\mathbb{C}$. For such functions ϕ, define $I_\mu(\phi)$ by*

$$I_\mu(\phi) = \sum_{y\in\mathbb{F}} y\mu(\phi^{-1}(\{y\})).$$

Proposition 19.12. *Let $\lambda \in \mathbb{F}$ and ϕ and ψ be two simple functions, then I_μ satisfies:*

1.
$$I_\mu(\lambda\phi) = \lambda I_\mu(\phi). \tag{19.8}$$

2.
$$I_\mu(\phi + \psi) = I_\mu(\psi) + I_\mu(\phi).$$

3. *If ϕ and ψ are non-negative simple functions such that $\phi \le \psi$ then*
$$I_\mu(\phi) \le I_\mu(\psi).$$

Proof. Let us write $\{\phi = y\}$ for the set $\phi^{-1}(\{y\}) \subset X$ and $\mu(\phi = y)$ for $\mu(\{\phi = y\}) = \mu(\phi^{-1}(\{y\}))$ so that

$$I_\mu(\phi) = \sum_{y\in\mathbb{F}} y\mu(\phi = y).$$

We will also write $\{\phi = a, \psi = b\}$ for $\phi^{-1}(\{a\}) \cap \psi^{-1}(\{b\})$. This notation is more intuitive for the purposes of this proof. Suppose that $\lambda \in \mathbb{F}$ then

$$\begin{aligned} I_\mu(\lambda\phi) &= \sum_{y\in\mathbb{F}} y\ \mu(\lambda\phi = y) = \sum_{y\in\mathbb{F}} y\ \mu(\phi = y/\lambda) \\ &= \sum_{z\in\mathbb{F}} \lambda z\ \mu(\phi = z) = \lambda I_\mu(\phi) \end{aligned}$$

provided that $\lambda \neq 0$. The case $\lambda = 0$ is clear, so we have proved 1. Suppose that ϕ and ψ are two simple functions, then

$$\begin{aligned} I_\mu(\phi + \psi) &= \sum_{z\in\mathbb{F}} z\ \mu(\phi + \psi = z) \\ &= \sum_{z\in\mathbb{F}} z\ \mu\left(\cup_{w\in\mathbb{F}} \{\phi = w,\ \psi = z - w\}\right) \\ &= \sum_{z\in\mathbb{F}} z \sum_{w\in\mathbb{F}} \mu(\phi = w,\ \psi = z - w) \\ &= \sum_{z,w\in\mathbb{F}} (z + w)\mu(\phi = w,\ \psi = z) \\ &= \sum_{z\in\mathbb{F}} z\ \mu(\psi = z) + \sum_{w\in\mathbb{F}} w\ \mu(\phi = w) \\ &= I_\mu(\psi) + I_\mu(\phi). \end{aligned}$$

which proves 2. For 3. if ϕ and ψ are non-negative simple functions such that $\phi \le \psi$

$$\begin{aligned} I_\mu(\phi) &= \sum_{a\geq 0} a\mu(\phi = a) = \sum_{a,b\geq 0} a\mu(\phi = a, \psi = b) \\ &\leq \sum_{a,b\geq 0} b\mu(\phi = a, \psi = b) = \sum_{b\geq 0} b\mu(\psi = b) = I_\mu(\psi), \end{aligned}$$

wherein the third inequality we have used $\{\phi = a, \psi = b\} = \emptyset$ if $a > b$. ■

19.3 Integrals of positive functions

Definition 19.13. *Let $L^+ = L^+(\mathcal{M}) = \{f : X \to [0, \infty] : f$ is measurable$\}$. Define*

$$\int_X f(x)\, d\mu(x) = \int_X f d\mu := \sup\{I_\mu(\phi) : \phi \text{ is simple and } \phi \leq f\}.$$

We say the $f \in L^+$ is ***integrable*** *if $\int_X f d\mu < \infty$. If $A \in \mathcal{M}$, let*

$$\int_A f(x)\, d\mu(x) = \int_A f d\mu := \int_X 1_A f \; d\mu.$$

Remark 19.14. Because of item 3. of Proposition 19.12, if ϕ is a non-negative simple function, $\int_X \phi d\mu = I_\mu(\phi)$ so that $\int_X$ is an extension of I_μ. This extension still has the monotonicity property if I_μ : namely if $0 \leq f \leq g$ then

$$\begin{aligned} \int_X f d\mu &= \sup\{I_\mu(\phi) : \phi \text{ is simple and } \phi \leq f\} \\ &\leq \sup\{I_\mu(\phi) : \phi \text{ is simple and } \phi \leq g\} \leq \int_X g d\mu. \end{aligned}$$

Similarly if $c > 0$,

$$\int_X c f d\mu = c \int_X f d\mu.$$

Also notice that if f is integrable, then $\mu(\{f = \infty\}) = 0$.

Lemma 19.15 (Sums as Integrals). *Let X be a set and $\rho : X \to [0, \infty]$ be a function, let $\mu = \sum_{x\in X} \rho(x)\delta_x$ on $\mathcal{M} = 2^X$, i.e.*

$$\mu(A) = \sum_{x\in A} \rho(x).$$

If $f : X \to [0, \infty]$ is a function (which is necessarily measurable), then

$$\int_X f d\mu = \sum_X f\rho.$$

Proof. Suppose that $\phi : X \to [0,\infty)$ is a simple function, then $\phi = \sum_{z\in[0,\infty)} z1_{\{\phi=z\}}$ and

$$\sum_X \phi\rho = \sum_{x\in X}\rho(x)\sum_{z\in[0,\infty)} z1_{\{\phi=z\}}(x) = \sum_{z\in[0,\infty)} z\sum_{x\in X}\rho(x)1_{\{\phi=z\}}(x)$$
$$= \sum_{z\in[0,\infty)} z\mu(\{\phi=z\}) = \int_X \phi d\mu.$$

So if $\phi : X \to [0,\infty)$ is a simple function such that $\phi \le f$, then

$$\int_X \phi d\mu = \sum_X \phi\rho \le \sum_X f\rho.$$

Taking the sup over ϕ in this last equation then shows that

$$\int_X f d\mu \le \sum_X f\rho.$$

For the reverse inequality, let $\Lambda \subset\subset X$ be a finite set and $N \in (0,\infty)$. Set $f^N(x) = \min\{N, f(x)\}$ and let $\phi_{N,\Lambda}$ be the simple function given by $\phi_{N,\Lambda}(x) := 1_\Lambda(x)f^N(x)$. Because $\phi_{N,\Lambda}(x) \le f(x)$,

$$\sum_\Lambda f^N\rho = \sum_X \phi_{N,\Lambda}\rho = \int_X \phi_{N,\Lambda} d\mu \le \int_X f d\mu.$$

Since $f^N \uparrow f$ as $N \to \infty$, we may let $N \to \infty$ in this last equation to concluded

$$\sum_\Lambda f\rho \le \int_X f d\mu.$$

Since Λ is arbitrary, this implies

$$\sum_X f\rho \le \int_X f d\mu.$$

■

Theorem 19.16 (Monotone Convergence Theorem). *Suppose $f_n \in L^+$ is a sequence of functions such that $f_n \uparrow f$ (f is necessarily in L^+) then*

$$\int f_n \uparrow \int f \text{ as } n \to \infty.$$

Proof. Since $f_n \le f_m \le f$, for all $n \le m < \infty$,

$$\int f_n \le \int f_m \le \int f$$

from which if follows $\int f_n$ is increasing in n and

$$\lim_{n\to\infty} \int f_n \leq \int f. \tag{19.9}$$

For the opposite inequality, let $\phi : X \to [0,\infty)$ be a simple function such that $0 \leq \phi \leq f$, $\alpha \in (0,1)$ and $X_n := \{f_n \geq \alpha\phi\}$. Notice that $X_n \uparrow X$ and $f_n \geq \alpha 1_{X_n}\phi$ and so by definition of $\int f_n$,

$$\int f_n \geq \int \alpha 1_{X_n}\phi = \alpha \int 1_{X_n}\phi. \tag{19.10}$$

Then using the continuity property of μ,

$$\begin{aligned}
\lim_{n\to\infty} \int 1_{X_n}\phi &= \lim_{n\to\infty} \int 1_{X_n} \sum_{y>0} y 1_{\{\phi=y\}} \\
&= \lim_{n\to\infty} \sum_{y>0} y\mu(X_n \cap \{\phi = y\}) = \sum_{y>0} y \lim_{n\to\infty} \mu(X_n \cap \{\phi = y\}) \\
&= \sum_{y>0} y \lim_{n\to\infty} \mu(\{\phi = y\}) = \int \phi.
\end{aligned}$$

This identity allows us to let $n \to \infty$ in Eq. (19.10) to conclude

$$\int_X \phi \leq \frac{1}{\alpha} \lim_{n\to\infty} \int f_n.$$

Since this is true for all non-negative simple functions ϕ with $\phi \leq f$;

$$\int f = \sup \left\{ \int_X \phi : \phi \text{ is simple and } \phi \leq f \right\} \leq \frac{1}{\alpha} \lim_{n\to\infty} \int f_n.$$

Because $\alpha \in (0,1)$ was arbitrary, it follows that $\int f \leq \lim_{n\to\infty} \int f_n$ which combined with Eq. (19.9) proves the theorem. ■

The following simple lemma will be use often in the sequel.

Lemma 19.17 (Chebyshev's Inequality). *Suppose that $f \geq 0$ is a measurable function, then for any $\varepsilon > 0$,*

$$\mu(f \geq \varepsilon) \leq \frac{1}{\varepsilon} \int_X f d\mu. \tag{19.11}$$

In particular if $\int_X f d\mu < \infty$ then $\mu(f = \infty) = 0$ (i.e. $f < \infty$ a.e.) and the set $\{f > 0\}$ is σ – finite.

Proof. Since $1_{\{f\geq\varepsilon\}} \leq 1_{\{f\geq\varepsilon\}} \frac{1}{\varepsilon} f \leq \frac{1}{\varepsilon} f$,

$$\mu(f \geq \varepsilon) = \int_X 1_{\{f\geq\varepsilon\}} d\mu \leq \int_X 1_{\{f\geq\varepsilon\}} \frac{1}{\varepsilon} f d\mu \leq \frac{1}{\varepsilon} \int_X f d\mu.$$

If $M := \int_X f d\mu < \infty$, then

$$\mu(f = \infty) \le \mu(f \ge n) \le \frac{M}{n} \to 0 \text{ as } n \to \infty$$

and $\{f \ge 1/n\} \uparrow \{f > 0\}$ with $\mu(f \ge 1/n) \le nM < \infty$ for all n. ∎

Corollary 19.18. *If $f_n \in L^+$ is a sequence of functions then*

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

In particular, if $\sum_{n=1}^{\infty} \int f_n < \infty$ then $\sum_{n=1}^{\infty} f_n < \infty$ a.e.

Proof. First off we show that

$$\int (f_1 + f_2) = \int f_1 + \int f_2$$

by choosing non-negative simple function ϕ_n and ψ_n such that $\phi_n \uparrow f_1$ and $\psi_n \uparrow f_2$. Then $(\phi_n + \psi_n)$ is simple as well and $(\phi_n + \psi_n) \uparrow (f_1 + f_2)$ so by the monotone convergence theorem,

$$\begin{aligned}\int (f_1 + f_2) &= \lim_{n\to\infty} \int (\phi_n + \psi_n) = \lim_{n\to\infty} \left(\int \phi_n + \int \psi_n \right) \\ &= \lim_{n\to\infty} \int \phi_n + \lim_{n\to\infty} \int \psi_n = \int f_1 + \int f_2.\end{aligned}$$

Now to the general case. Let $g_N := \sum\limits_{n=1}^{N} f_n$ and $g = \sum\limits_{1}^{\infty} f_n$, then $g_N \uparrow g$ and so again by monotone convergence theorem and the additivity just proved,

$$\begin{aligned}\sum_{n=1}^{\infty} \int f_n &:= \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n = \lim_{N\to\infty} \int \sum_{n=1}^{N} f_n \\ &= \lim_{N\to\infty} \int g_N = \int g =: \int \sum_{n=1}^{\infty} f_n.\end{aligned}$$

∎

Remark 19.19. It is in the proof of this corollary (i.e. the linearity of the integral) that we really make use of the assumption that all of our functions are measurable. In fact the definition $\int f d\mu$ makes sense for **all** functions $f : X \to [0, \infty]$ not just measurable functions. Moreover the monotone convergence theorem holds in this generality with no change in the proof. However, in the proof of Corollary 19.18, we use the approximation Theorem 18.42 which relies heavily on the measurability of the functions to be approximated.

The following Lemma and the next Corollary are simple applications of Corollary 19.18.

Lemma 19.20 (The First Borell – Carntelli Lemma). *Let $(X, \mathcal{M}, \mu)$ be a measure space, $A_n \in \mathcal{M}$, and set*

$$\{A_n \text{ i.o.}\} = \{x \in X : x \in A_n \text{ for infinitely many } n\text{'s}\} = \bigcap_{N=1}^{\infty} \bigcup_{n \geq N} A_n.$$

If $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ then $\mu(\{A_n \text{ i.o.}\}) = 0$.

Proof. (First Proof.) Let us first observe that

$$\{A_n \text{ i.o.}\} = \left\{ x \in X : \sum_{n=1}^{\infty} 1_{A_n}(x) = \infty \right\}.$$

Hence if $\sum_{n=1}^{\infty} \mu(A_n) < \infty$ then

$$\infty > \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \int_X 1_{A_n} \, d\mu = \int_X \sum_{n=1}^{\infty} 1_{A_n} \, d\mu$$

implies that $\sum\limits_{n=1}^{\infty} 1_{A_n}(x) < \infty$ for μ - a.e. x. That is to say $\mu(\{A_n \text{ i.o.}\}) = 0$.
(Second Proof.) Of course we may give a strictly measure theoretic proof of this fact:

$$\begin{aligned} \mu(A_n \text{ i.o.}) &= \lim_{N \to \infty} \mu \left(\bigcup_{n \geq N} A_n \right) \\ &\leq \lim_{N \to \infty} \sum_{n \geq N} \mu(A_n) \end{aligned}$$

and the last limit is zero since $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. ∎

Corollary 19.21. *Suppose that $(X, \mathcal{M}, \mu)$ is a measure space and $\{A_n\}_{n=1}^{\infty} \subset \mathcal{M}$ is a collection of sets such that $\mu(A_i \cap A_j) = 0$ for all $i \neq j$, then*

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

Proof. Since

$$\begin{aligned} \mu\left(\cup_{n=1}^{\infty} A_n\right) &= \int_X 1_{\cup_{n=1}^{\infty} A_n} d\mu \text{ and} \\ \sum_{n=1}^{\infty} \mu(A_n) &= \int_X \sum_{n=1}^{\infty} 1_{A_n} d\mu \end{aligned}$$

it suffices to show

$$\sum_{n=1}^{\infty} 1_{A_n} = 1_{\cup_{n=1}^{\infty} A_n} \ \mu - \text{a.e.} \tag{19.12}$$

Now $\sum_{n=1}^{\infty} 1_{A_n} \geq 1_{\cup_{n=1}^{\infty} A_n}$ and $\sum_{n=1}^{\infty} 1_{A_n}(x) \neq 1_{\cup_{n=1}^{\infty} A_n}(x)$ iff $x \in A_i \cap A_j$ for some $i \neq j$, that is

$$\left\{ x : \sum_{n=1}^{\infty} 1_{A_n}(x) \neq 1_{\cup_{n=1}^{\infty} A_n}(x) \right\} = \cup_{i<j} A_i \cap A_j$$

and the latter set has measure 0 being the countable union of sets of measure zero. This proves Eq. (19.12) and hence the corollary. ■

Notation 19.22 *If m is Lebesgue measure on $\mathcal{B}_{\mathbb{R}}$, f is a non-negative Borel measurable function and $a < b$ with $a, b \in \bar{\mathbb{R}}$, we will often write $\int_a^b f(x)\,dx$ or $\int_a^b f dm$ for $\int_{(a,b]\cap\mathbb{R}} f dm$.*

Example 19.23. Suppose $-\infty < a < b < \infty$, $f \in C([a,b],[0,\infty))$ and m be Lebesgue measure on $\mathbb{R}$. Also let $\pi_k = \{a = a_0^k < a_1^k < \cdots < a_{n_k}^k = b\}$ be a sequence of refining partitions (i.e. $\pi_k \subset \pi_{k+1}$ for all k) such that

$$\text{mesh}(\pi_k) := \max\{\left|a_j^k - a_{j-1}^{k+1}\right| : j = 1, \ldots, n_k\} \to 0 \text{ as } k \to \infty.$$

For each k, let

$$f_k(x) = f(a) 1_{\{a\}} + \sum_{l=0}^{n_k-1} \min\left\{f(x) : a_l^k \leq x \leq a_{l+1}^k\right\} 1_{(a_l^k, a_{l+1}^k]}(x)$$

then $f_k \uparrow f$ as $k \to \infty$ and so by the monotone convergence theorem,

$$\begin{aligned}
\int_a^b f dm &:= \int_{[a,b]} f dm = \lim_{k\to\infty} \int_a^b f_k \ dm \\
&= \lim_{k\to\infty} \sum_{l=0}^{n_k-1} \min\left\{f(x) : a_l^k \leq x \leq a_{l+1}^k\right\} m\left((a_l^k, a_{l+1}^k]\right) \\
&= \int_a^b f(x) dx.
\end{aligned}$$

The latter integral being the Riemann integral.

We can use the above result to integrate some non-Riemann integrable functions:

Example 19.24. For all $\lambda > 0$,

$$\int_0^{\infty} e^{-\lambda x} dm(x) = \lambda^{-1} \text{ and } \int_{\mathbb{R}} \frac{1}{1+x^2} dm(x) = \pi.$$

The proof of these identities are similar. By the monotone convergence theorem, Example 19.23 and the fundamental theorem of calculus for Riemann integrals (or see Theorem 8.13 above or Theorem 19.40 below),

$$\begin{aligned}\int_0^\infty e^{-\lambda x} dm(x) &= \lim_{N\to\infty}\int_0^N e^{-\lambda x} dm(x) = \lim_{N\to\infty}\int_0^N e^{-\lambda x} dx \\ &= -\lim_{N\to\infty}\frac{1}{\lambda}e^{-\lambda x}|_0^N = \lambda^{-1}\end{aligned}$$

and

$$\begin{aligned}\int_{\mathbb{R}}\frac{1}{1+x^2}dm(x) &= \lim_{N\to\infty}\int_{-N}^N\frac{1}{1+x^2}dm(x) = \lim_{N\to\infty}\int_{-N}^N\frac{1}{1+x^2}dx \\ &= \lim_{N\to\infty}\left[\tan^{-1}(N)-\tan^{-1}(-N)\right] = \pi.\end{aligned}$$

Let us also consider the functions x^{-p},

$$\begin{aligned}\int_{(0,1]}\frac{1}{x^p}\,dm(x) &= \lim_{n\to\infty}\int_0^1 1_{(\frac{1}{n},1]}(x)\frac{1}{x^p}dm(x) \\ &= \lim_{n\to\infty}\int_{\frac{1}{n}}^1\frac{1}{x^p}dx = \lim_{n\to\infty}\left.\frac{x^{-p+1}}{1-p}\right|_{1/n}^1 \\ &= \begin{cases}\frac{1}{1-p} & \text{if } p<1 \\ \infty & \text{if } p>1\end{cases}\end{aligned}$$

If $p=1$ we find

$$\int_{(0,1]}\frac{1}{x^p}\,dm(x) = \lim_{n\to\infty}\int_{\frac{1}{n}}^1\frac{1}{x}dx = \lim_{n\to\infty}\ln(x)|_{1/n}^1 = \infty.$$

Example 19.25. Let $\{r_n\}_{n=1}^\infty$ be an enumeration of the points in $\mathbb{Q}\cap[0,1]$ and define

$$f(x) = \sum_{n=1}^\infty 2^{-n}\frac{1}{\sqrt{|x-r_n|}}$$

with the convention that

$$\frac{1}{\sqrt{|x-r_n|}} = 5 \text{ if } x = r_n.$$

Since, By Theorem 19.40,

$$\begin{aligned}\int_0^1\frac{1}{\sqrt{|x-r_n|}}\,dx &= \int_{r_n}^1\frac{1}{\sqrt{x-r_n}}dx + \int_0^{r_n}\frac{1}{\sqrt{r_n-x}}dx \\ &= 2\sqrt{x-r_n}|_{r_n}^1 - 2\sqrt{r_n-x}|_0^{r_n} = 2\left(\sqrt{1-r_n}-\sqrt{r_n}\right) \\ &\le 4,\end{aligned}$$

we find

$$\int_{[0,1]} f(x)dm(x) = \sum_{n=1}^{\infty} 2^{-n} \int_{[0,1]} \frac{1}{\sqrt{|x-r_n|}} dx \leq \sum_{n=1}^{\infty} 2^{-n} 4 = 4 < \infty.$$

In particular, $m(f = \infty) = 0$, i.e. that $f < \infty$ for almost every $x \in [0,1]$ and this implies that

$$\sum_{n=1}^{\infty} 2^{-n} \frac{1}{\sqrt{|x-r_n|}} < \infty \text{ for a.e. } x \in [0,1].$$

This result is somewhat surprising since the singularities of the summands form a dense subset of $[0,1]$.

Proposition 19.26. *Suppose that $f \geq 0$ is a measurable function. Then $\int_X f d\mu = 0$ iff $f = 0$ a.e. Also if $f, g \geq 0$ are measurable functions such that $f \leq g$ a.e. then $\int f d\mu \leq \int g d\mu$. In particular if $f = g$ a.e. then $\int f d\mu = \int g d\mu$.*

Proof. If $f = 0$ a.e. and $\phi \leq f$ is a simple function then $\phi = 0$ a.e. This implies that $\mu(\phi^{-1}(\{y\})) = 0$ for all $y > 0$ and hence $\int_X \phi d\mu = 0$ and therefore $\int_X f d\mu = 0$. Conversely, if $\int f d\mu = 0$, then by Chebyshev's Inequality (Lemma 19.17),

$$\mu(f \geq 1/n) \leq n \int f d\mu = 0 \text{ for all } n.$$

Therefore, $\mu(f > 0) \leq \sum_{n=1}^{\infty} \mu(f \geq 1/n) = 0$, i.e. $f = 0$ a.e. For the second assertion let E be the exceptional set where $f > g$, i.e. $E := \{x \in X : f(x) > g(x)\}$. By assumption E is a null set and $1_{E^c} f \leq 1_{E^c} g$ everywhere. Because $g = 1_{E^c} g + 1_E g$ and $1_E g = 0$ a.e.,

$$\int g d\mu = \int 1_{E^c} g d\mu + \int 1_E g d\mu = \int 1_{E^c} g d\mu$$

and similarly $\int f d\mu = \int 1_{E^c} f d\mu$. Since $1_{E^c} f \leq 1_{E^c} g$ everywhere,

$$\int f d\mu = \int 1_{E^c} f d\mu \leq \int 1_{E^c} g d\mu = \int g d\mu.$$

∎

Corollary 19.27. *Suppose that $\{f_n\}$ is a sequence of non-negative measurable functions and f is a measurable function such that $f_n \uparrow f$ off a null set, then*

$$\int f_n \uparrow \int f \text{ as } n \to \infty.$$

Proof. Let $E \subset X$ be a null set such that $f_n 1_{E^c} \uparrow f 1_{E^c}$ as $n \to \infty$. Then by the monotone convergence theorem and Proposition 19.26,

$$\int f_n = \int f_n 1_{E^c} \uparrow \int f 1_{E^c} = \int f \text{ as } n \to \infty.$$

■

Lemma 19.28 (Fatou's Lemma). *If $f_n : X \to [0, \infty]$ is a sequence of measurable functions then*

$$\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n$$

Proof. Define $g_k := \inf_{n\ge k} f_n$ so that $g_k \uparrow \liminf_{n\to\infty} f_n$ as $k \to \infty$. Since $g_k \le f_n$ for all $k \le n$,

$$\int g_k \le \int f_n \text{ for all } n \ge k$$

and therefore

$$\int g_k \le \lim \inf_{n\to\infty} \int f_n \text{ for all } k.$$

We may now use the monotone convergence theorem to let $k \to \infty$ to find

$$\int \lim \inf_{n\to\infty} f_n = \int \lim_{k\to\infty} g_k \overset{\text{MCT}}{=} \lim_{k\to\infty} \int g_k \le \lim \inf_{n\to\infty} \int f_n.$$

■

19.4 Integrals of Complex Valued Functions

Definition 19.29. *A measurable function $f : X \to \bar{\mathbb{R}}$ is **integrable** if $f_+ := f 1_{\{f\ge 0\}}$ and $f_- = -f\ 1_{\{f\le 0\}}$ are **integrable**. We write $\mathrm{L}^1(\mu;\mathbb{R})$ for the space of real valued integrable functions. For $f \in \mathrm{L}^1(\mu;\mathbb{R})$, let*

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

Convention: If $f, g : X \to \bar{\mathbb{R}}$ are two measurable functions, let $f + g$ denote the collection of measurable functions $h : X \to \bar{\mathbb{R}}$ such that $h(x) = f(x)+g(x)$ whenever $f(x)+g(x)$ is well defined, i.e. is not of the form $\infty-\infty$ or $-\infty+\infty$. We use a similar convention for $f - g$. Notice that if $f, g \in \mathrm{L}^1(\mu;\mathbb{R})$ and $h_1, h_2 \in f + g$, then $h_1 = h_2$ a.e. because $|f| < \infty$ and $|g| < \infty$ a.e.

Notation 19.30 (Abuse of notation) *We will sometimes denote the integral $\int_X f d\mu$ by $\mu(f)$. With this notation we have $\mu(A) = \mu(1_A)$ for all $A \in \mathcal{M}$.*

Remark 19.31. Since
$$f_\pm \le |f| \le f_+ + f_-,$$
a measurable function f is **integrable** iff $\int |f|\ d\mu < \infty$. Hence

$$\mathrm{L}^1(\mu;\mathbb{R}) := \left\{ f : X \to \bar{\mathbb{R}} :\ f \text{ is measurable and } \int_X |f|\ d\mu < \infty \right\}.$$

If $f, g \in \mathrm{L}^1(\mu;\mathbb{R})$ and $f = g$ a.e. then $f_\pm = g_\pm$ a.e. and so it follows from Proposition 19.26 that $\int f d\mu = \int g d\mu$. In particular if $f, g \in \mathrm{L}^1(\mu;\mathbb{R})$ we may define
$$\int_X (f+g)\, d\mu = \int_X h d\mu$$
where h is any element of $f+g$.

Proposition 19.32. *The map*

$$f \in \mathrm{L}^1(\mu;\mathbb{R}) \to \int_X f d\mu \in \mathbb{R}$$

is linear and has the monotonicity property: $\int f d\mu \le \int g d\mu$ for all $f, g \in \mathrm{L}^1(\mu;\mathbb{R})$ such that $f \le g$ a.e.

Proof. Let $f, g \in \mathrm{L}^1(\mu;\mathbb{R})$ and $a, b \in \mathbb{R}$. By modifying f and g on a null set, we may assume that f, g are real valued functions. We have $af + bg \in \mathrm{L}^1(\mu;\mathbb{R})$ because

$$|af + bg| \le |a|\,|f| + |b|\,|g| \in \mathrm{L}^1(\mu;\mathbb{R}).$$

If $a < 0$, then
$$(af)_+ = -af_- \text{ and } (af)_- = -af_+$$
so that

$$\int af = -a \int f_- + a \int f_+ = a(\int f_+ - \int f_-) = a \int f.$$

A similar calculation works for $a > 0$ and the case $a = 0$ is trivial so we have shown that
$$\int af = a \int f.$$
Now set $h = f + g$. Since $h = h_+ - h_-$,

$$h_+ - h_- = f_+ - f_- + g_+ - g_-$$

or

$$h_+ + f_- + g_- = h_- + f_+ + g_+.$$

Therefore,

$$\int h_+ + \int f_- + \int g_- = \int h_- + \int f_+ + \int g_+$$

and hence

$$\int h = \int h_+ - \int h_- = \int f_+ + \int g_+ - \int f_- - \int g_- = \int f + \int g.$$

Finally if $f_+ - f_- = f \leq g = g_+ - g_-$ then $f_+ + g_- \leq g_+ + f_-$ which implies that

$$\int f_+ + \int g_- \leq \int g_+ + \int f_-$$

or equivalently that

$$\int f = \int f_+ - \int f_- \leq \int g_+ - \int g_- = \int g.$$

The monotonicity property is also a consequence of the linearity of the integral, the fact that $f \leq g$ a.e. implies $0 \leq g - f$ a.e. and Proposition 19.26. ■

Definition 19.33. *A measurable function $f : X \to \mathbb{C}$ is **integrable** if $\int_X |f| \ d\mu < \infty$. Analogously to the real case, let*

$$\mathrm{L}^1(\mu;\mathbb{C}) := \left\{ f : X \to \mathbb{C} : \ f \text{ is measurable and } \int_X |f| \ d\mu < \infty \right\}.$$

denote the complex valued integrable functions. Because, $\max(|\mathrm{Re}\, f|, |\mathrm{Im}\, f|) \leq |f| \leq \sqrt{2} \max(|\mathrm{Re}\, f|, |\mathrm{Im}\, f|)$, $\int |f| \ d\mu < \infty$ *iff*

$$\int |\mathrm{Re}\, f| \, d\mu + \int |\mathrm{Im}\, f| \, d\mu < \infty.$$

For $f \in \mathrm{L}^1(\mu;\mathbb{C})$ define

$$\int f \ d\mu = \int \mathrm{Re}\, f \ d\mu + i \int \mathrm{Im}\, f \ d\mu.$$

It is routine to show the integral is still linear on $\mathrm{L}^1(\mu;\mathbb{C})$ (prove!). In the remainder of this section, let $\mathrm{L}^1(\mu)$ be either $\mathrm{L}^1(\mu;\mathbb{C})$ or $\mathrm{L}^1(\mu;\mathbb{R})$. If $A \in \mathcal{M}$ and $f \in \mathrm{L}^1(\mu;\mathbb{C})$ or $f : X \to [0,\infty]$ is a measurable function, let

$$\int_A f d\mu := \int_X 1_A f d\mu.$$

Proposition 19.34. *Suppose that $f \in \mathrm{L}^1(\mu;\mathbb{C})$, then*

$$\left| \int_X f d\mu \right| \leq \int_X |f| \, d\mu. \tag{19.13}$$

Proof. Start by writing $\int_X f\, d\mu = Re^{i\theta}$ with $R \geq 0$. We may assume that $R = \left|\int_X f d\mu\right| > 0$ since otherwise there is nothing to prove. Since

$$R = e^{-i\theta} \int_X f\, d\mu = \int_X e^{-i\theta} f\, d\mu = \int_X \mathrm{Re}\left(e^{-i\theta} f\right) d\mu + i \int_X \mathrm{Im}\left(e^{-i\theta} f\right) d\mu,$$

it must be that $\int_X \mathrm{Im}\left[e^{-i\theta} f\right] d\mu = 0$. Using the monotonicity in Proposition 19.26,

$$\left|\int_X f d\mu\right| = \int_X \mathrm{Re}\left(e^{-i\theta} f\right) d\mu \leq \int_X \left|\mathrm{Re}\left(e^{-i\theta} f\right)\right| d\mu \leq \int_X |f|\, d\mu.$$

■

Proposition 19.35. *Let $f, g \in \mathrm{L}^1(\mu)$, then*

1. *The set $\{f \neq 0\}$ is σ – finite, in fact $\{|f| \geq \frac{1}{n}\} \uparrow \{f \neq 0\}$ and $\mu(|f| \geq \frac{1}{n}) < \infty$ for all n.*
2. *The following are equivalent*
   a) $\int_E f = \int_E g$ for all $E \in \mathcal{M}$
   b) $\int\limits_X |f - g| = 0$
   c) $f = g$ a.e.

Proof. 1. By Chebyshev's inequality, Lemma 19.17,

$$\mu(|f| \geq \frac{1}{n}) \leq n \int_X |f|\, d\mu < \infty$$

for all n. 2. (a) $\Longrightarrow$ (c) Notice that

$$\int_E f = \int_E g \Leftrightarrow \int_E (f - g) = 0$$

for all $E \in \mathcal{M}$. Taking $E = \{\mathrm{Re}(f-g) > 0\}$ and using $1_E \mathrm{Re}(f-g) \geq 0$, we learn that

$$0 = \mathrm{Re} \int_E (f-g) d\mu = \int 1_E \mathrm{Re}(f-g) \Longrightarrow 1_E \mathrm{Re}(f-g) = 0 \text{ a.e.}$$

This implies that $1_E = 0$ a.e. which happens iff

$$\mu\left(\{\mathrm{Re}(f-g) > 0\}\right) = \mu(E) = 0.$$

Similar $\mu(\mathrm{Re}(f-g) < 0) = 0$ so that $\mathrm{Re}(f-g) = 0$ a.e. Similarly, $\mathrm{Im}(f-g) = 0$ a.e and hence $f - g = 0$ a.e., i.e. $f = g$ a.e. (c) $\Longrightarrow$ (b) is clear and so is (b) $\Longrightarrow$ (a) since

$$\left|\int_E f - \int_E g\right| \leq \int |f - g| = 0.$$

■

Definition 19.36. *Let $(X, \mathcal{M}, \mu)$ be a measure space and $L^1(\mu) = L^1(X, \mathcal{M}, \mu)$ denote the set of $\mathrm{L}^1(\mu)$ functions modulo the equivalence relation; $f \sim g$ iff $f = g$ a.e. We make this into a normed space using the norm*

$$\|f - g\|_{L^1} = \int |f - g| \, d\mu$$

and into a metric space using $\rho_1(f, g) = \|f - g\|_{L^1}$.

Warning: in the future we will often not make much of a distinction between $L^1(\mu)$ and $\mathrm{L}^1(\mu)$. On occasion this can be dangerous and this danger will be pointed out when necessary.

Remark 19.37. More generally we may define $L^p(\mu) = L^p(X, \mathcal{M}, \mu)$ for $p \in [1, \infty)$ as the set of measurable functions f such that

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

modulo the equivalence relation; $f \sim g$ iff $f = g$ a.e.

We will see in Chapter 21 that

$$\|f\|_{L^p} = \left(\int |f|^p \, d\mu \right)^{1/p} \quad \text{for } f \in L^p(\mu)$$

is a norm and $(L^p(\mu), \|\cdot\|_{L^p})$ is a Banach space in this norm.

Theorem 19.38 (Dominated Convergence Theorem). *Suppose $f_n, g_n, g \in \mathrm{L}^1(\mu)$, $f_n \to f$ a.e., $|f_n| \le g_n \in \mathrm{L}^1(\mu)$, $g_n \to g$ a.e. and $\int_X g_n d\mu \to \int_X g d\mu$. Then $f \in \mathrm{L}^1(\mu)$ and*

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

(In most typical applications of this theorem $g_n = g \in \mathrm{L}^1(\mu)$ for all n.)

Proof. Notice that $|f| = \lim_{n \to \infty} |f_n| \le \lim_{n \to \infty} |g_n| \le g$ a.e. so that $f \in \mathrm{L}^1(\mu)$. By considering the real and imaginary parts of f separately, it suffices to prove the theorem in the case where f is real. By Fatou's Lemma,

$$\begin{aligned}
\int_X (g \pm f) d\mu &= \int_X \liminf_{n \to \infty} (g_n \pm f_n) \, d\mu \le \liminf_{n \to \infty} \int_X (g_n \pm f_n) \, d\mu \\
&= \lim_{n \to \infty} \int_X g_n d\mu + \liminf_{n \to \infty} \left(\pm \int_X f_n d\mu \right) \\
&= \int_X g d\mu + \liminf_{n \to \infty} \left(\pm \int_X f_n d\mu \right)
\end{aligned}$$

Since $\liminf_{n \to \infty}(-a_n) = -\limsup_{n \to \infty} a_n$, we have shown,

$$\int_X g d\mu \pm \int_X f d\mu \leq \int_X g d\mu + \begin{cases} \liminf_{n\to\infty} \int_X f_n d\mu \\ -\limsup_{n\to\infty} \int_X f_n d\mu \end{cases}$$

and therefore

$$\limsup_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

This shows that $\lim_{n\to\infty} \int_X f_n d\mu$ exists and is equal to $\int_X f d\mu$. ∎

Exercise 19.1. Give another proof of Proposition 19.34 by first proving Eq. (19.13) with f being a cylinder function in which case the triangle inequality for complex numbers will do the trick. Then use the approximation Theorem 18.42 along with the dominated convergence Theorem 19.38 to handle the general case.

Corollary 19.39. *Let $\{f_n\}_{n=1}^\infty \subset \mathrm{L}^1(\mu)$ be a sequence such that $\sum_{n=1}^\infty \|f_n\|_{\mathrm{L}^1(\mu)} < \infty$, then $\sum_{n=1}^\infty f_n$ is convergent a.e. and*

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

Proof. The condition $\sum_{n=1}^\infty \|f_n\|_{\mathrm{L}^1(\mu)} < \infty$ is equivalent to $\sum_{n=1}^\infty |f_n| \in \mathrm{L}^1(\mu)$. Hence $\sum_{n=1}^\infty f_n$ is almost everywhere convergent and if $S_N := \sum_{n=1}^N f_n$, then

$$|S_N| \leq \sum_{n=1}^N |f_n| \leq \sum_{n=1}^\infty |f_n| \in \mathrm{L}^1(\mu).$$

So by the dominated convergence theorem,

$$\begin{aligned}\int_X \left(\sum_{n=1}^\infty f_n\right) d\mu &= \int_X \lim_{N\to\infty} S_N d\mu = \lim_{N\to\infty} \int_X S_N d\mu \\ &= \lim_{N\to\infty} \sum_{n=1}^N \int_X f_n d\mu = \sum_{n=1}^\infty \int_X f_n d\mu.\end{aligned}$$

∎

Theorem 19.40 (The Fundamental Theorem of Calculus). *Suppose $-\infty < a < b < \infty$, $f \in C((a,b),\mathbb{R}) \cap L^1((a,b),m)$ and $F(x) := \int_a^x f(y)dm(y)$. Then*

1. *$F \in C([a,b],\mathbb{R}) \cap C^1((a,b),\mathbb{R})$.*
2. *$F'(x) = f(x)$ for all $x \in (a,b)$.*
3. *If $G \in C([a,b],\mathbb{R}) \cap C^1((a,b),\mathbb{R})$ is an anti-derivative of f on (a,b) (i.e. $f = G'|_{(a,b)}$) then*

$$\int_a^b f(x)dm(x) = G(b) - G(a).$$

Proof. Since $F(x) := \int_{\mathbb{R}} 1_{(a,x)}(y)f(y)dm(y)$, $\lim_{x\to z} 1_{(a,x)}(y) = 1_{(a,z)}(y)$ for m – a.e. y and $\left|1_{(a,x)}(y)f(y)\right| \leq 1_{(a,b)}(y)\left|f(y)\right|$ is an L^1 – function, it follows from the dominated convergence Theorem 19.38 that F is continuous on $[a,b]$. Simple manipulations show,

$$\begin{aligned}
\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| &= \frac{1}{|h|}\begin{cases} \left|\int_x^{x+h}\left[f(y)-f(x)\right]dm(y)\right| & \text{if } h>0 \\ \left|\int_{x+h}^{x}\left[f(y)-f(x)\right]dm(y)\right| & \text{if } h<0 \end{cases} \\
&\leq \frac{1}{|h|}\begin{cases} \int_x^{x+h}\left|f(y)-f(x)\right|dm(y) \text{ if } h>0 \\ \int_{x+h}^{x}\left|f(y)-f(x)\right|dm(y) \text{ if } h<0 \end{cases} \\
&\leq \sup\left\{\left|f(y)-f(x)\right| : y \in \left[x-|h|, x+|h|\right]\right\}
\end{aligned}$$

and the latter expression, by the continuity of f, goes to zero as $h \to 0$. This shows $F' = f$ on (a,b). For the converse direction, we have by assumption that $G'(x) = F'(x)$ for $x \in (a,b)$. Therefore by the mean value theorem, $F - G = C$ for some constant C. Hence

$$\begin{aligned}
\int_a^b f(x)dm(x) &= F(b) = F(b) - F(a) \\
&= (G(b)+C) - (G(a)+C) = G(b) - G(a).
\end{aligned}$$

■

Example 19.41. The following limit holds,

$$\lim_{n\to\infty}\int_0^n (1-\frac{x}{n})^n dm(x) = 1.$$

Let $f_n(x) = (1-\frac{x}{n})^n 1_{[0,n]}(x)$ and notice that $\lim_{n\to\infty} f_n(x) = e^{-x}$. We will now show

$$0 \leq f_n(x) \leq e^{-x} \text{ for all } x \geq 0.$$

It suffices to consider $x \in [0,n]$. Let $g(x) = e^x f_n(x)$, then for $x \in (0,n)$,

$$\frac{d}{dx}\ln g(x) = 1 + n\frac{1}{(1-\frac{x}{n})}(-\frac{1}{n}) = 1 - \frac{1}{(1-\frac{x}{n})} \leq 0$$

which shows that $\ln g(x)$ and hence $g(x)$ is decreasing on $[0,n]$. Therefore $g(x) \leq g(0) = 1$, i.e.

$$0 \leq f_n(x) \leq e^{-x}.$$

From Example 19.24, we know

$$\int_0^\infty e^{-x}dm(x) = 1 < \infty,$$

so that e^{-x} is an integrable function on $[0,\infty)$. Hence by the dominated convergence theorem,

$$\lim_{n\to\infty}\int_0^n (1-\frac{x}{n})^n dm(x) = \lim_{n\to\infty}\int_0^\infty f_n(x)dm(x)$$
$$= \int_0^\infty \lim_{n\to\infty} f_n(x)dm(x) = \int_0^\infty e^{-x}dm(x) = 1.$$

Example 19.42 (Integration of Power Series). Suppose $R>0$ and $\{a_n\}_{n=0}^\infty$ is a sequence of complex numbers such that $\sum_{n=0}^\infty |a_n|\, r^n < \infty$ for all $r\in(0,R)$. Then

$$\int_\alpha^\beta \left(\sum_{n=0}^\infty a_n x^n\right) dm(x) = \sum_{n=0}^\infty a_n \int_\alpha^\beta x^n dm(x) = \sum_{n=0}^\infty a_n \frac{\beta^{n+1}-\alpha^{n+1}}{n+1}$$

for all $-R<\alpha<\beta<R$. Indeed this follows from Corollary 19.39 since

$$\sum_{n=0}^\infty \int_\alpha^\beta |a_n|\,|x|^n\, dm(x) \le \sum_{n=0}^\infty \left(\int_0^{|\beta|} |a_n|\,|x|^n\, dm(x) + \int_0^{|\alpha|} |a_n|\,|x|^n\, dm(x)\right)$$
$$\le \sum_{n=0}^\infty |a_n| \frac{|\beta|^{n+1}+|\alpha|^{n+1}}{n+1} \le 2r\sum_{n=0}^\infty |a_n|\, r^n < \infty$$

where $r=\max(|\beta|\,,|\alpha|)$.

Corollary 19.43 (Differentiation Under the Integral). *Suppose that $J\subset\mathbb{R}$ is an open interval and $f: J\times X\to\mathbb{C}$ is a function such that*

1. *$x\to f(t,x)$ is measurable for each $t\in J$.*
2. *$f(t_0,\cdot)\in L^1(\mu)$ for some $t_0\in J$.*
3. *$\frac{\partial f}{\partial t}(t,x)$ exists for all (t,x).*
4. *There is a function $g\in \mathrm{L}^1(\mu)$ such that $\left|\frac{\partial f}{\partial t}(t,\cdot)\right| \le g\in\mathrm{L}^1(\mu)$ for each $t\in J$.*

Then $f(t,\cdot)\in\mathrm{L}^1(\mu)$ for all $t\in J$ (i.e. $\int_X |f(t,x)|\,d\mu(x)<\infty$), $t\to \int_X f(t,x)d\mu(x)$ is a differentiable function on J and

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

Proof. (The proof is essentially the same as for sums.) By considering the real and imaginary parts of f separately, we may assume that f is real. Also notice that

$$\frac{\partial f}{\partial t}(t,x) = \lim_{n\to\infty} n(f(t+n^{-1},x)-f(t,x))$$

and therefore, for $x\to\frac{\partial f}{\partial t}(t,x)$ is a sequential limit of measurable functions and hence is measurable for all $t\in J$. By the mean value theorem,

$$|f(t,x)-f(t_0,x)| \le g(x)\,|t-t_0| \text{ for all } t\in J \tag{19.14}$$

and hence

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

This shows $f(t,\cdot) \in \mathrm{L}^1(\mu)$ for all $t \in J$. Let $G(t) := \int_X f(t,x)d\mu(x)$, then

$$\frac{G(t)-G(t_0)}{t-t_0} = \int_X \frac{f(t,x)-f(t_0,x)}{t-t_0} d\mu(x).$$

By assumption,

$$\lim_{t\to t_0} \frac{f(t,x)-f(t_0,x)}{t-t_0} = \frac{\partial f}{\partial t}(t,x) \text{ for all } x \in X$$

and by Eq. (19.14),

$$\left| \frac{f(t,x)-f(t_0,x)}{t-t_0} \right| \le g(x) \text{ for all } t \in J \text{ and } x \in X.$$

Therefore, we may apply the dominated convergence theorem to conclude

$$\begin{aligned} \lim_{n\to\infty} \frac{G(t_n)-G(t_0)}{t_n-t_0} &= \lim_{n\to\infty} \int_X \frac{f(t_n,x)-f(t_0,x)}{t_n-t_0} d\mu(x) \\ &= \int_X \lim_{n\to\infty} \frac{f(t_n,x)-f(t_0,x)}{t_n-t_0} d\mu(x) \\ &= \int_X \frac{\partial f}{\partial t}(t_0,x) d\mu(x) \end{aligned}$$

for **all** sequences $t_n \in J \setminus \{t_0\}$ such that $t_n \to t_0$. Therefore, $\dot{G}(t_0) = \lim_{t\to t_0} \frac{G(t)-G(t_0)}{t-t_0}$ exists and

$$\dot{G}(t_0) = \int_X \frac{\partial f}{\partial t}(t_0,x) d\mu(x).$$

■

Example 19.44. Recall from Example 19.24 that

$$\lambda^{-1} = \int_{[0,\infty)} e^{-\lambda x} dm(x) \text{ for all } \lambda > 0.$$

Let $\varepsilon > 0$. For $\lambda \ge 2\varepsilon > 0$ and $n \in \mathbb{N}$ there exists $C_n(\varepsilon) < \infty$ such that

$$0 \le \left(-\frac{d}{d\lambda}\right)^n e^{-\lambda x} = x^n e^{-\lambda x} \le C(\varepsilon) e^{-\varepsilon x}.$$

Using this fact, Corollary 19.43 and induction gives

$$\begin{aligned} n!\lambda^{-n-1} &= \left(-\frac{d}{d\lambda}\right)^n \lambda^{-1} = \int_{[0,\infty)} \left(-\frac{d}{d\lambda}\right)^n e^{-\lambda x} dm(x) \\ &= \int_{[0,\infty)} x^n e^{-\lambda x} dm(x). \end{aligned}$$

That is $n! = \lambda^n \int_{[0,\infty)} x^n e^{-\lambda x} dm(x)$. Recall that

$$\Gamma(t) := \int_{[0,\infty)} x^{t-1} e^{-x} dx \text{ for } t > 0.$$

(The reader should check that $\Gamma(t) < \infty$ for all $t > 0$.) We have just shown that $\Gamma(n+1) = n!$ for all $n \in \mathbb{N}$.

Remark 19.45. Corollary 19.43 may be generalized by allowing the hypothesis to hold for $x \in X \setminus E$ where $E \in \mathcal{M}$ is a **fixed** null set, i.e. E must be independent of t. Consider what happens if we formally apply Corollary 19.43 to $g(t) := \int_0^\infty 1_{x \le t} dm(x)$,

$$\dot{g}(t) = \frac{d}{dt} \int_0^\infty 1_{x \le t} dm(x) \stackrel{?}{=} \int_0^\infty \frac{\partial}{\partial t} 1_{x \le t} dm(x).$$

The last integral is zero since $\frac{\partial}{\partial t} 1_{x \le t} = 0$ unless $t = x$ in which case it is not defined. On the other hand $g(t) = t$ so that $\dot{g}(t) = 1$. (The reader should decide which hypothesis of Corollary 19.43 has been violated in this example.)

19.5 Measurability on Complete Measure Spaces

In this subsection we will discuss a couple of measurability results concerning completions of measure spaces.

Proposition 19.46. *Suppose that $(X, \mathcal{M}, \mu)$ is a complete measure space*[1] *and $f : X \to \mathbb{R}$ is measurable.*

1. *If $g : X \to \mathbb{R}$ is a function such that $f(x) = g(x)$ for μ – a.e. x, then g is measurable.*
2. *If $f_n : X \to \mathbb{R}$ are measurable and $f : X \to \mathbb{R}$ is a function such that $\lim_{n\to\infty} f_n = f$, μ - a.e., then f is measurable as well.*

Proof. 1. Let $E = \{x : f(x) \ne g(x)\}$ which is assumed to be in $\mathcal{M}$ and $\mu(E) = 0$. Then $g = 1_{E^c} f + 1_E g$ since $f = g$ on E^c. Now $1_{E^c} f$ is measurable so g will be measurable if we show $1_E g$ is measurable. For this consider,

$$(1_E g)^{-1}(A) = \begin{cases} E^c \cup (1_E g)^{-1}(A \setminus \{0\}) & \text{if } 0 \in A \\ (1_E g)^{-1}(A) & \text{if } 0 \notin A \end{cases} \tag{19.15}$$

Since $(1_E g)^{-1}(B) \subset E$ if $0 \notin B$ and $\mu(E) = 0$, it follow by completeness of $\mathcal{M}$ that $(1_E g)^{-1}(B) \in \mathcal{M}$ if $0 \notin B$. Therefore Eq. (19.15) shows that $1_E g$ is measurable. 2. Let $E = \{x : \lim_{n\to\infty} f_n(x) \ne f(x)\}$ by assumption $E \in \mathcal{M}$ and

[1] Recall this means that if $N \subset X$ is a set such that $N \subset A \in \mathcal{M}$ and $\mu(A) = 0$, then $N \in \mathcal{M}$ as well.

$\mu(E) = 0$. Since $g := 1_E f = \lim_{n\to\infty} 1_{E^c} f_n$, g is measurable. Because $f = g$ on E^c and $\mu(E) = 0$, $f = g$ a.e. so by part 1. f is also measurable. ■

The above results are in general false if $(X, \mathcal{M}, \mu)$ is not complete. For example, let $X = \{0, 1, 2\}$, $\mathcal{M} = \{\{0\}, \{1, 2\}, X, \phi\}$ and $\mu = \delta_0$. Take $g(0) = 0$, $g(1) = 1$, $g(2) = 2$, then $g = 0$ a.e. yet g is not measurable.

Lemma 19.47. *Suppose that $(X, \mathcal{M}, \mu)$ is a measure space and $\bar{\mathcal{M}}$ is the completion of $\mathcal{M}$ relative to μ and $\bar{\mu}$ is the extension of μ to $\bar{\mathcal{M}}$. Then a function $f : X \to \mathbb{R}$ is $(\bar{\mathcal{M}}, \mathcal{B} = \mathcal{B}_{\mathbb{R}})$ – measurable iff there exists a function $g : X \to \mathbb{R}$ that is $(\mathcal{M}, \mathcal{B})$ – measurable such $E = \{x : f(x) \neq g(x)\} \in \bar{\mathcal{M}}$ and $\bar{\mu}(E) = 0$, i.e. $f(x) = g(x)$ for $\bar{\mu}$ – a.e. x. Moreover for such a pair f and g, $f \in L^1(\bar{\mu})$ iff $g \in L^1(\mu)$ and in which case*

$$\int_X f d\bar{\mu} = \int_X g d\mu.$$

Proof. Suppose first that such a function g exists so that $\bar{\mu}(E) = 0$. Since g is also $(\bar{\mathcal{M}}, \mathcal{B})$ – measurable, we see from Proposition 19.46 that f is $(\bar{\mathcal{M}}, \mathcal{B})$ – measurable. Conversely if f is $(\bar{\mathcal{M}}, \mathcal{B})$ – measurable, by considering $f_\pm$ we may assume that $f \geq 0$. Choose $(\bar{\mathcal{M}}, \mathcal{B})$ – measurable simple function $\phi_n \geq 0$ such that $\phi_n \uparrow f$ as $n \to \infty$. Writing

$$\phi_n = \sum a_k 1_{A_k}$$

with $A_k \in \bar{\mathcal{M}}$, we may choose $B_k \in \mathcal{M}$ such that $B_k \subset A_k$ and $\bar{\mu}(A_k \backslash B_k) = 0$. Letting

$$\tilde{\phi}_n := \sum a_k 1_{B_k}$$

we have produced a $(\mathcal{M}, \mathcal{B})$ – measurable simple function $\tilde{\phi}_n \geq 0$ such that $E_n := \{\phi_n \neq \tilde{\phi}_n\}$ has zero $\bar{\mu}$ – measure. Since $\bar{\mu}(\cup_n E_n) \leq \sum_n \bar{\mu}(E_n)$, there exists $F \in \mathcal{M}$ such that $\cup_n E_n \subset F$ and $\mu(F) = 0$. It now follows that

$$1_F \tilde{\phi}_n = 1_F \phi_n \uparrow g := 1_F f \text{ as } n \to \infty.$$

This shows that $g = 1_F f$ is $(\mathcal{M}, \mathcal{B})$ – measurable and that $\{f \neq g\} \subset F$ has $\bar{\mu}$ – measure zero. Since $f = g$, $\bar{\mu}$ – a.e., $\int_X f d\bar{\mu} = \int_X g d\bar{\mu}$ so to prove Eq. (19.16) it suffices to prove

$$\int_X g d\bar{\mu} = \int_X g d\mu. \tag{19.16}$$

Because $\bar{\mu} = \mu$ on $\mathcal{M}$, Eq. (19.16) is easily verified for non-negative $\mathcal{M}$ – measurable simple functions. Then by the monotone convergence theorem and the approximation Theorem 18.42 it holds for all $\mathcal{M}$ – measurable functions $g : X \to [0, \infty]$. The rest of the assertions follow in the standard way by considering $(\operatorname{Re} g)_\pm$ and $(\operatorname{Im} g)_\pm$. ■

19.6 Comparison of the Lebesgue and the Riemann Integral

For the rest of this chapter, let $-\infty < a < b < \infty$ and $f : [a,b] \to \mathbb{R}$ be a bounded function. A partition of $[a,b]$ is a finite subset $\pi \subset [a,b]$ containing $\{a,b\}$. To each partition

$$\pi = \{a = t_0 < t_1 < \cdots < t_n = b\} \tag{19.17}$$

of $[a,b]$ let

$$\operatorname{mesh}(\pi) := \max\{|t_j - t_{j-1}| : j = 1, \ldots, n\},$$

$$M_j = \sup\{f(x) : t_j \le x \le t_{j-1}\}, \quad m_j = \inf\{f(x) : t_j \le x \le t_{j-1}\}$$
$$G_\pi = f(a)1_{\{a\}} + \sum_1^n M_j 1_{(t_{j-1},t_j]}, \quad g_\pi = f(a)1_{\{a\}} + \sum_1^n m_j 1_{(t_{j-1},t_j]} \text{ and}$$

$$S_\pi f = \sum M_j(t_j - t_{j-1}) \text{ and } s_\pi f = \sum m_j(t_j - t_{j-1}).$$

Notice that

$$S_\pi f = \int_a^b G_\pi dm \text{ and } s_\pi f = \int_a^b g_\pi dm.$$

The upper and lower Riemann integrals are defined respectively by

$$\overline{\int_a^b} f(x)dx = \inf_\pi S_\pi f \text{ and } \underline{\int_b^a} f(x)dx = \sup_\pi \ s_\pi f.$$

Definition 19.48. *The function f is* ***Riemann integrable*** *iff $\overline{\int_a^b} f = \underline{\int_a^b} f \in \mathbb{R}$ and which case the Riemann integral $\int_a^b f$ is defined to be the common value:*

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

The proof of the following Lemma is left to the reader as Exercise 19.20.

Lemma 19.49. *If π' and π are two partitions of $[a,b]$ and $\pi \subset \pi'$ then*

$$\begin{aligned} &G_\pi \ge G_{\pi'} \ge f \ge g_{\pi'} \ge g_\pi \text{ and} \\ &S_\pi f \ge S_{\pi'} f \ge s_{\pi'} f \ge s_\pi f. \end{aligned}$$

There exists an increasing sequence of partitions $\{\pi_k\}_{k=1}^\infty$ such that $\operatorname{mesh}(\pi_k) \downarrow 0$ *and*

$$S_{\pi_k} f \downarrow \overline{\int_a^b} f \text{ and } s_{\pi_k} f \uparrow \underline{\int_a^b} f \text{ as } k \to \infty.$$

If we let

$$G := \lim_{k\to\infty} G_{\pi_k} \text{ and } g := \lim_{k\to\infty} g_{\pi_k} \tag{19.18}$$

then by the dominated convergence theorem,

$$\int_{[a,b]} g dm = \lim_{k\to\infty} \int_{[a,b]} g_{\pi_k} = \lim_{k\to\infty} s_{\pi_k} f = \underline{\int_a^b} f(x) dx \tag{19.19}$$

and

$$\int_{[a,b]} G dm = \lim_{k\to\infty} \int_{[a,b]} G_{\pi_k} = \lim_{k\to\infty} S_{\pi_k} f = \overline{\int_a^b} f(x) dx. \tag{19.20}$$

Notation 19.50 *For $x \in [a,b]$, let*

$$H(x) = \limsup_{y\to x} f(y) := \lim_{\varepsilon\downarrow 0} \sup\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\} \text{ and}$$
$$h(x) = \liminf_{y\to x} f(y) := \lim_{\varepsilon\downarrow 0} \inf\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\}.$$

Lemma 19.51. *The functions $H, h : [a,b] \to \mathbb{R}$ satisfy:*

1. *$h(x) \le f(x) \le H(x)$ for all $x \in [a,b]$ and $h(x) = H(x)$ iff f is continuous at x.*
2. *If $\{\pi_k\}_{k=1}^\infty$ is any increasing sequence of partitions such that $\operatorname{mesh}(\pi_k) \downarrow 0$ and G and g are defined as in Eq. (19.18), then*
$$G(x) = H(x) \ge f(x) \ge h(x) = g(x) \quad \forall\ x \notin \pi := \cup_{k=1}^\infty \pi_k. \tag{19.21}$$
(Note π is a countable set.)
3. *H and h are Borel measurable.*

Proof. Let $G_k := G_{\pi_k} \downarrow G$ and $g_k := g_{\pi_k} \uparrow g$.

1. It is clear that $h(x) \le f(x) \le H(x)$ for all x and $H(x) = h(x)$ iff $\lim_{y\to x} f(y)$ exists and is equal to $f(x)$. That is $H(x) = h(x)$ iff f is continuous at x.
2. For $x \notin \pi$,
$$G_k(x) \ge H(x) \ge f(x) \ge h(x) \ge g_k(x)\ \forall\ k$$
and letting $k \to \infty$ in this equation implies
$$G(x) \ge H(x) \ge f(x) \ge h(x) \ge g(x)\ \forall\ x \notin \pi. \tag{19.22}$$
Moreover, given $\varepsilon > 0$ and $x \notin \pi$,
$$\sup\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\} \ge G_k(x)$$
for all k large enough, since eventually $G_k(x)$ is the supremum of $f(y)$ over some interval contained in $[x-\varepsilon, x+\varepsilon]$. Again letting $k \to \infty$ implies $\sup_{|y-x|\le\varepsilon} f(y) \ge G(x)$ and therefore, that

$$H(x) = \limsup_{y \to x} f(y) \geq G(x)$$

for all $x \notin \pi$. Combining this equation with Eq. (19.22) then implies $H(x) = G(x)$ if $x \notin \pi$. A similar argument shows that $h(x) = g(x)$ if $x \notin \pi$ and hence Eq. (19.21) is proved.

3. The functions G and g are limits of measurable functions and hence measurable. Since $H = G$ and $h = g$ except possibly on the countable set π, both H and h are also Borel measurable. (You justify this statement.)

■

Theorem 19.52. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then*

$$\overline{\int_a^b} f = \int_{[a,b]} H dm \text{ and } \underline{\int_a^b} f = \int_{[a,b]} h dm \tag{19.23}$$

and the following statements are equivalent:

1. *$H(x) = h(x)$ for m -a.e. x,*
2. *the set*
$$E := \{x \in [a, b] : f \text{ is discontinuous at } x\}$$
is an $\bar{m}$ – null set.
3. *f is Riemann integrable.*

If f is Riemann integrable then f is Lebesgue measurable[2]*, i.e. f is $\mathcal{L}/\mathcal{B}$ – measurable where $\mathcal{L}$ is the Lebesgue σ – algebra and $\mathcal{B}$ is the Borel σ – algebra on $[a, b]$. Moreover if we let $\bar{m}$ denote the completion of m, then*

$$\int_{[a,b]} H dm = \int_a^b f(x)dx = \int_{[a,b]} f d\bar{m} = \int_{[a,b]} h dm. \tag{19.24}$$

Proof. Let $\{\pi_k\}_{k=1}^\infty$ be an increasing sequence of partitions of $[a, b]$ as described in Lemma 19.49 and let G and g be defined as in Lemma 19.51. Since $m(\pi) = 0$, $H = G$ a.e., Eq. (19.23) is a consequence of Eqs. (19.19) and (19.20). From Eq. (19.23), f is Riemann integrable iff

$$\int_{[a,b]} H dm = \int_{[a,b]} h dm$$

and because $h \leq f \leq H$ this happens iff $h(x) = H(x)$ for m - a.e. x. Since $E = \{x : H(x) \neq h(x)\}$, this last condition is equivalent to E being a m – null set. In light of these results and Eq. (19.21), the remaining assertions including Eq. (19.24) are now consequences of Lemma 19.47. ■

Notation 19.53 *In view of this theorem we will often write $\int_a^b f(x)dx$ for $\int_a^b f dm$.*

---

[2] f need not be Borel measurable.

19.7 Determining Classes of Measures

Definition 19.54 (σ – finite). *Let X be a set and $\mathcal{E} \subset \mathcal{F} \subset 2^X$. We say that a function $\mu : \mathcal{F} \to [0, \infty]$ is σ –* ***finite on*** *$\mathcal{E}$ if there exist $X_n \in \mathcal{E}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ for all n.*

Theorem 19.55 (Uniqueness). *Suppose that $\mathcal{C} \subset 2^X$ is a π – class (see Definition 18.53), $\mathcal{M} = \sigma(\mathcal{C})$ and μ and ν are two measure on $\mathcal{M}$. If μ and ν are σ – finite on $\mathcal{C}$ and $\mu = \nu$ on $\mathcal{C}$, then $\mu = \nu$ on $\mathcal{M}$.*

Proof. We begin first with the special case where $\mu(X) < \infty$ and therefore also

$$\nu(X) = \lim_{n\to\infty} \nu(X_n) = \lim_{n\to\infty} \mu(X_n) = \mu(X) < \infty.$$

Let

$$\mathcal{H} := \{f \in \ell^\infty(\mathcal{M}, \mathbb{R}) : \mu(f) = \nu(f)\}.$$

Then $\mathcal{H}$ is a linear subspace which is closed under bounded convergence (by the dominated convergence theorem), contains 1 and contains the multiplicative system, $M := \{1_C : C \in \mathcal{C}\}$. Therefore, by Theorem 18.51 or Corollary 18.54, $\mathcal{H} = \ell^\infty(\mathcal{M}, \mathbb{R})$ and hence $\mu = \nu$. For the general case, let $X_n^1, X_n^2 \in \mathcal{C}$ be chosen so that $X_n^1 \uparrow X$ and $X_n^2 \uparrow X$ as $n \to \infty$ and $\mu\left(X_n^1\right) + \nu\left(X_n^2\right) < \infty$ for all n. Then $X_n := X_n^1 \cap X_n^2 \in \mathcal{C}$ increases to X and $\nu(X_n) = \mu(X_n) < \infty$ for all n. For each $n \in \mathbb{N}$, define two measures μ_n and ν_n on $\mathcal{M}$ by

$$\mu_n(A) := \mu(A \cap X_n) \text{ and } \nu_n(A) = \nu(A \cap X_n).$$

Then, as the reader should verify, μ_n and ν_n are finite measure on $\mathcal{M}$ such that $\mu_n = \nu_n$ on $\mathcal{C}$. Therefore, by the special case just proved, $\mu_n = \nu_n$ on $\mathcal{M}$. Finally, using the continuity properties of measures,

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap X_n) = \lim_{n\to\infty} \nu(A \cap X_n) = \nu(A)$$

for all $A \in \mathcal{M}$. ■

As an immediate consequence we have the following corollaries.

Corollary 19.56. *Suppose that (X, τ) is a topological space, $\mathcal{B}_X = \sigma(\tau)$ is the Borel σ – algebra on X and μ and ν are two measures on $\mathcal{B}_X$ which are σ – finite on τ. If $\mu = \nu$ on τ then $\mu = \nu$ on $\mathcal{B}_X$, i.e. $\mu \equiv \nu$.*

Corollary 19.57. *Suppose that μ and ν are two measures on $\mathcal{B}_{\mathbb{R}^n}$ which are finite on bounded sets and such that $\mu(A) = \nu(A)$ for all sets A of the form*

$$A = (a, b] = (a_1, b_1] \times \cdots \times (a_n, b_n]$$

with $a, b \in \mathbb{R}^n$ and $a < b$, i.e. $a_i < b_i$ for all i. Then $\mu = \nu$ on $\mathcal{B}_{\mathbb{R}^n}$.

Proposition 19.58. *Suppose that (X,d) is a metric space, μ and ν are two measures on $\mathcal{B}_X := \sigma(\tau_d)$ which are finite on bounded measurable subsets of X and*

$$\int_X f d\mu = \int_X f d\nu \tag{19.25}$$

for all $f \in BC_b(X,\mathbb{R})$ where

$$BC_b(X,\mathbb{R}) = \{f \in BC(X,\mathbb{R}) : \operatorname{supp}(f) \text{ is bounded}\}. \tag{19.26}$$

Then $\mu \equiv \nu$.

Proof. To prove this fix a $o \in X$ and let

$$\psi_R(x) = ([R+1-d(x,o)] \wedge 1) \vee 0 \tag{19.27}$$

so that $\psi_R \in BC_b(X,[0,1])$, $\operatorname{supp}(\psi_R) \subset B(o,R+2)$ and $\psi_R \uparrow 1$ as $R \to \infty$. Let $\mathcal{H}_R$ denote the space of bounded real valued $\mathcal{B}_X$ – measurable functions f such that

$$\int_X \psi_R f d\mu = \int_X \psi_R f d\nu. \tag{19.28}$$

Then $\mathcal{H}_R$ is closed under bounded convergence and because of Eq. (19.25) contains $BC(X,\mathbb{R})$. Therefore by Corollary 18.55, $\mathcal{H}_R$ contains all bounded measurable functions on X. Take $f = 1_A$ in Eq. (19.28) with $A \in \mathcal{B}_X$, and then use the monotone convergence theorem to let $R \to \infty$. The result is $\mu(A) = \nu(A)$ for all $A \in \mathcal{B}_X$. ■

Here is another version of Proposition 19.58.

Proposition 19.59. *Suppose that (X,d) is a metric space, μ and ν are two measures on $\mathcal{B}_X = \sigma(\tau_d)$ which are both finite on compact sets. Further assume there exists compact sets $K_k \subset X$ such that $K_k^o \uparrow X$. If*

$$\int_X f d\mu = \int_X f d\nu \tag{19.29}$$

for all $f \in C_c(X,\mathbb{R})$ then $\mu \equiv \nu$.

Proof. Let $\psi_{n,k}$ be defined as in the proof of Proposition 18.56 and let $\mathcal{H}_{n,k}$ denote those bounded $\mathcal{B}_X$ – measurable functions, $f : X \to \mathbb{R}$ such that

$$\int_X f\psi_{n,k} d\mu = \int_X f\psi_{n,k} d\nu.$$

By assumption $BC(X,\mathbb{R}) \subset \mathcal{H}_{n,k}$ and one easily checks that $\mathcal{H}_{n,k}$ is closed under bounded convergence. Therefore, by Corollary 18.55, $\mathcal{H}_{n,k}$ contains all bounded measurable function. In particular for $A \in \mathcal{B}_X$,

$$\int_X 1_A \psi_{n,k} d\mu = \int_X 1_A \psi_{n,k} d\nu.$$

Letting $n \to \infty$ in this equation, using the dominated convergence theorem, one shows

$$\int_X 1_A 1_{K_k^o} d\mu = \int_X 1_A 1_{K_k^o} d\nu$$

holds for k. Finally using the monotone convergence theorem we may let $k \to \infty$ to conclude

$$\mu(A) = \int_X 1_A d\mu = \int_X 1_A d\nu = \nu(A)$$

for all $A \in \mathcal{B}_X$. ∎

19.8 Exercises

Exercise 19.2. Let μ be a measure on an algebra $\mathcal{A} \subset 2^X$, then $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$ for all $A, B \in \mathcal{A}$.

Exercise 19.3 (From problem 12 on p. 27 of Folland.). Let $(X, \mathcal{M}, \mu)$ be a finite measure space and for $A, B \in \mathcal{M}$ let $\rho(A, B) = \mu(A \Delta B)$ where $A\Delta B = (A \setminus B) \cup (B \setminus A)$. It is clear that $\rho(A, B) = \rho(B, A)$. Show:

1. ρ satisfies the triangle inequality:
$$\rho(A, C) \leq \rho(A, B) + \rho(B, C) \text{ for all } A, B, C \in \mathcal{M}.$$
2. Define $A \sim B$ iff $\mu(A\Delta B) = 0$ and notice that $\rho(A, B) = 0$ iff $A \sim B$. Show "$\sim$ " is an equivalence relation.
3. Let $\mathcal{M}/\sim$ denote $\mathcal{M}$ modulo the equivalence relation, $\sim$, and let $[A] := \{B \in \mathcal{M} : B \sim A\}$. Show that $\bar{\rho}([A], [B]) := \rho(A, B)$ is gives a well defined metric on $\mathcal{M}/\sim$.
4. Similarly show $\tilde{\mu}([A]) = \mu(A)$ is a well defined function on $\mathcal{M}/\sim$ and show $\tilde{\mu} : (\mathcal{M}/\sim) \to \mathbb{R}_+$ is $\bar{\rho}-$ continuous.

Exercise 19.4. Suppose that $\mu_n : \mathcal{M} \to [0, \infty]$ are measures on $\mathcal{M}$ for $n \in \mathbb{N}$. Also suppose that $\mu_n(A)$ is increasing in n for all $A \in \mathcal{M}$. Prove that $\mu : \mathcal{M} \to [0, \infty]$ defined by $\mu(A) := \lim_{n\to\infty} \mu_n(A)$ is also a measure.

Exercise 19.5. Now suppose that Λ is some index set and for each $\lambda \in \Lambda$, $\mu_\lambda : \mathcal{M} \to [0, \infty]$ is a measure on $\mathcal{M}$. Define $\mu : \mathcal{M} \to [0, \infty]$ by $\mu(A) = \sum_{\lambda\in\Lambda} \mu_\lambda(A)$ for each $A \in \mathcal{M}$. Show that μ is also a measure.

Exercise 19.6. Let $(X, \mathcal{M}, \mu)$ be a measure space and $\rho : X \to [0, \infty]$ be a measurable function. For $A \in \mathcal{M}$, set $\nu(A) := \int_A \rho d\mu$.

1. Show $\nu : \mathcal{M} \to [0, \infty]$ is a measure.

2. Let $f : X \to [0,\infty]$ be a measurable function, show
$$\int_X f d\nu = \int_X f\rho d\mu. \tag{19.30}$$
Hint: first prove the relationship for characteristic functions, then for simple functions, and then for general positive measurable functions.
3. Show that a measurable function $f : X \to \mathbb{C}$ is in $L^1(\nu)$ iff $|f|\,\rho \in L^1(\mu)$ and if $f \in L^1(\nu)$ then Eq. (19.30) still holds.

Notation 19.60 *It is customary to informally describe ν defined in Exercise 19.6 by writing $d\nu = \rho d\mu$.*

Exercise 19.7. Let $(X, \mathcal{M}, \mu)$ be a measure space, $(Y, \mathcal{F})$ be a measurable space and $f : X \to Y$ be a measurable map. Define a function $\nu : \mathcal{F} \to [0,\infty]$ by $\nu(A) := \mu(f^{-1}(A))$ for all $A \in \mathcal{F}$.

1. Show ν is a measure. (We will write $\nu = f_*\mu$ or $\nu = \mu \circ f^{-1}$.)
2. Show
$$\int_Y g d\nu = \int_X (g \circ f)\, d\mu \tag{19.31}$$
for all measurable functions $g : Y \to [0,\infty]$. **Hint:** see the hint from Exercise 19.6.
3. Show a measurable function $g : Y \to \mathbb{C}$ is in $L^1(\nu)$ iff $g \circ f \in L^1(\mu)$ and that Eq. (19.31) holds for all $g \in \mathrm{L}^1(\nu)$.

Exercise 19.8. Let $F : \mathbb{R} \to \mathbb{R}$ be a C^1-function such that $F'(x) > 0$ for all $x \in \mathbb{R}$ and $\lim_{x\to\pm\infty} F(x) = \pm\infty$. (Notice that F is strictly increasing so that $F^{-1} : \mathbb{R} \to \mathbb{R}$ exists and moreover, by the inverse function theorem that F^{-1} is a C^1 – function.) Let m be Lebesgue measure on $\mathcal{B}_\mathbb{R}$ and
$$\nu(A) = m(F(A)) = m(\left(F^{-1}\right)^{-1}(A)) = \left(F_*^{-1}m\right)(A)$$
for all $A \in \mathcal{B}_\mathbb{R}$. Show $d\nu = F'dm$. Use this result to prove the change of variable formula,
$$\int_\mathbb{R} h \circ F \cdot F' dm = \int_\mathbb{R} h dm \tag{19.32}$$
which is valid for all Borel measurable functions $h : \mathbb{R} \to [0,\infty]$.

Hint: Start by showing $d\nu = F'dm$ on sets of the form $A = (a,b]$ with $a,b \in \mathbb{R}$ and $a < b$. Then use the uniqueness assertions in Theorem 19.8 (or see Corollary 19.57) to conclude $d\nu = F'dm$ on all of $\mathcal{B}_\mathbb{R}$. To prove Eq. (19.32) apply Exercise 19.7 with $g = h \circ F$ and $f = F^{-1}$.

Exercise 19.9. Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{A_n\}_{n=1}^\infty \subset \mathcal{M}$, show
$$\mu(\{A_n \text{ a.a.}\}) \le \liminf_{n\to\infty} \mu(A_n)$$
and if $\mu(\cup_{m\ge n} A_m) < \infty$ for some n, then
$$\mu(\{A_n \text{ i.o.}\}) \ge \limsup_{n\to\infty} \mu(A_n).$$

Exercise 19.10. BRUCE: Delete this exercise which is contained in Lemma 19.17. Suppose $(X, \mathcal{M}, \mu)$ be a measure space and $f : X \to [0\,\infty]$ be a measurable function such that $\int_X f d\mu < \infty$. Show $\mu(\{f = \infty\}) = 0$ and the set $\{f > 0\}$ is σ – finite.

Exercise 19.11. Folland 2.13 on p. 52. **Hint:** "Fatou times two."

Exercise 19.12. Folland 2.14 on p. 52. BRUCE: delete this exercise

Exercise 19.13. Give examples of measurable functions $\{f_n\}$ on $\mathbb{R}$ such that f_n decreases to 0 uniformly yet $\int f_n dm = \infty$ for all n. Also give an example of a sequence of measurable functions $\{g_n\}$ on $[0, 1]$ such that $g_n \to 0$ while $\int g_n dm = 1$ for all n.

Exercise 19.14. Folland 2.19 on p. 59. (This problem is essentially covered in the previous exercise.)

Exercise 19.15. Suppose $\{a_n\}_{n=-\infty}^{\infty} \subset \mathbb{C}$ is a summable sequence (i.e. $\sum_{n=-\infty}^{\infty} |a_n| < \infty$), then $f(\theta) := \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ is a continuous function for $\theta \in \mathbb{R}$ and

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta.$$

Exercise 19.16. For any function $f \in L^1(m)$, show $x \in \mathbb{R} \to \int_{(-\infty,x]} f(t)\, dm(t)$ is continuous in x. Also find a finite measure, μ, on $\mathcal{B}_{\mathbb{R}}$ such that $x \to \int_{(-\infty,x]} f(t)\, d\mu(t)$ is not continuous.

Exercise 19.17. Folland 2.28 on p. 60.

Exercise 19.18. Folland 2.31b and 2.31e on p. 60. (The answer in 2.13b is wrong by a factor of -1 and the sum is on $k = 1$ to ∞. In part e, s should be taken to be a. You may also freely use the Taylor series expansion

$$(1-z)^{-1/2} = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{2^n n!} z^n = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)^2} z^n \text{ for } |z| < 1.$$

Exercise 19.19. There exists a meager (see Definition 13.4 and Proposition 13.3) subsets of $\mathbb{R}$ which have full Lebesgue measure, i.e. whose complement is a Lebesgue null set. (This is Folland 5.27. **Hint:** Consider the generalized Cantor sets discussed on p. 39 of Folland.)

Exercise 19.20. Prove Lemma 19.49.

20

Multiple Integrals

In this chapter we will introduce iterated integrals and product measures. We are particularly interested in when it is permissible to interchange the order of integration in multiple integrals.

Example 20.1. As an example let $X = [1, \infty)$ and $Y = [0, 1]$ equipped with their Borel σ - algebras and let $\mu = \nu = m$, where m is Lebesgue measure. The iterated integrals of the function $f(x, y) := e^{-xy} - 2e^{-2xy}$ satisfy,

$$\int_0^1 \left[\int_1^\infty (e^{-xy} - 2e^{-2xy}) dx \right] dy = \int_0^1 e^{-y} \left(\frac{1 - e^{-y}}{y} \right) dy \in (0, \infty)$$

and

$$\int_1^\infty \left[\int_0^1 (e^{-xy} - 2e^{-2xy}) dy \right] dx = -\int_1^\infty e^{-x} \left[\frac{1 - e^{-x}}{x} \right] dx \in (-\infty, 0)$$

and therefore are not equal. Hence it is not always true that order of integration is irrelevant.

Lemma 20.2. *Let $\mathbb{F}$ be either $[0, \infty)$, $\mathbb{R}$ or $\mathbb{C}$. Suppose $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are two measurable spaces and $f : X \times Y \to \mathbb{F}$ is a $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{F}})$ – measurable function, then for each $y \in Y$,*

$$x \to f(x, y) \text{ is } (\mathcal{M}, \mathcal{B}_{\mathbb{F}}) \text{ measurable,} \tag{20.1}$$

for each $x \in X$,

$$y \to f(x, y) \text{ is } (\mathcal{N}, \mathcal{B}_{\mathbb{F}}) \text{ measurable.} \tag{20.2}$$

Proof. Suppose that $E = A \times B \in \mathcal{E} := \mathcal{M} \times \mathcal{N}$ and $f = 1_E$. Then

$$f(x, y) = 1_{A \times B}(x, y) = 1_A(x) 1_B(y)$$

from which it follows that Eqs. (20.1) and (20.2) for this function. Let $\mathcal{H}$ be the collection of all bounded $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{F}})$ – measurable functions on $X \times Y$ such that Eqs. (20.1) and (20.2) hold, here we assume $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Because measurable functions are closed under taking linear combinations and pointwise limits, $\mathcal{H}$ is linear subspace of $\ell^\infty(\mathcal{M}\otimes\mathcal{N},\mathbb{F})$ which is closed under bounded convergence and contain $1_E\in\mathcal{H}$ for all E in the π – class, $\mathcal{E}$. Therefore by by Corollary 18.54, that $\mathcal{H}=\ell^\infty(\mathcal{M}\otimes\mathcal{N},\mathbb{F})$.

For the general $(\mathcal{M}\otimes\mathcal{N},\mathcal{B}_{\mathbb{R}})$ – measurable functions $f:X\times Y\to\mathbb{F}$ and $M\in\mathbb{N}$, let $f_M:=1_{|f|\le M}f\in\ell^\infty(\mathcal{M}\otimes\mathcal{N},\mathbb{F})$. Then Eqs. (20.1) and (20.2) hold with f replaced by f_M and hence for f as well by letting $M\to\infty$. ■

Notation 20.3 (Iterated Integrals) *If $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are two measure spaces and $f:X\times Y\to\mathbb{C}$ is a $\mathcal{M}\otimes\mathcal{N}$ – measurable function, the* ***iterated integrals*** *of f (when they make sense) are:*

$$\int_X d\mu(x)\int_Y d\nu(y)f(x,y):=\int_X\left[\int_Y f(x,y)d\nu(y)\right]d\mu(x)$$

and

$$\int_Y d\nu(y)\int_X d\mu(x)f(x,y):=\int_Y\left[\int_X f(x,y)d\mu(x)\right]d\nu(y).$$

Notation 20.4 *Suppose that $f:X\to\mathbb{C}$ and $g:Y\to\mathbb{C}$ are functions, let $f\otimes g$ denote the function on $X\times Y$ given by*

$$f\otimes g(x,y)=f(x)g(y).$$

Notice that if f,g are measurable, then $f\otimes g$ is $(\mathcal{M}\otimes\mathcal{N},\mathcal{B}_{\mathbb{C}})$ – measurable. To prove this let $F(x,y)=f(x)$ and $G(x,y)=g(y)$ so that $f\otimes g=F\cdot G$ will be measurable provided that F and G are measurable. Now $F=f\circ\pi_1$ where $\pi_1:X\times Y\to X$ is the projection map. This shows that F is the composition of measurable functions and hence measurable. Similarly one shows that G is measurable.

20.1 Fubini-Tonelli's Theorem and Product Measure

Theorem 20.5. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces and f is a nonnegative $(\mathcal{M}\otimes\mathcal{N},\mathcal{B}_{\mathbb{R}})$ – measurable function, then for each $y\in Y$,*

$$x\to f(x,y)\text{ is }\mathcal{M}-\mathcal{B}_{[0,\infty]}\text{ measurable,}\tag{20.3}$$

for each $x\in X$,

$$y\to f(x,y)\text{ is }\mathcal{N}-\mathcal{B}_{[0,\infty]}\text{ measurable,}\tag{20.4}$$

$$x\to\int_Y f(x,y)d\nu(y)\text{ is }\mathcal{M}-\mathcal{B}_{[0,\infty]}\text{ measurable,}\tag{20.5}$$

$$y\to\int_X f(x,y)d\mu(x)\text{ is }\mathcal{N}-\mathcal{B}_{[0,\infty]}\text{ measurable,}\tag{20.6}$$

and

$$\int_X d\mu(x) \int_Y d\nu(y) f(x,y) = \int_Y d\nu(y) \int_X d\mu(x) f(x,y). \tag{20.7}$$

Proof. Suppose that $E = A \times B \in \mathcal{E} := \mathcal{M} \times \mathcal{N}$ and $f = 1_E$. Then

$$f(x,y) = 1_{A\times B}(x,y) = 1_A(x)1_B(y)$$

and one sees that Eqs. (20.3) and (20.4) hold. Moreover

$$\int_Y f(x,y)d\nu(y) = \int_Y 1_A(x)1_B(y)d\nu(y) = 1_A(x)\nu(B),$$

so that Eq. (20.5) holds and we have

$$\int_X d\mu(x) \int_Y d\nu(y) f(x,y) = \nu(B)\mu(A). \tag{20.8}$$

Similarly,

$$\int_X f(x,y)d\mu(x) = \mu(A)1_B(y) \text{ and}$$
$$\int_Y d\nu(y) \int_X d\mu(x) f(x,y) = \nu(B)\mu(A)$$

from which it follows that Eqs. (20.6) and (20.7) hold in this case as well. For the moment let us further assume that $\mu(X) < \infty$ and $\nu(Y) < \infty$ and let $\mathcal{H}$ be the collection of all bounded $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X \times Y$ such that Eqs. (20.3) – (20.7) hold. Using the fact that measurable functions are closed under pointwise limits and the dominated convergence theorem (the dominating function always being a constant), one easily shows that $\mathcal{H}$ closed under bounded convergence. Since we have just verified that $1_E \in \mathcal{H}$ for all E in the π – class, $\mathcal{E}$, it follows by Corollary 18.54 that $\mathcal{H}$ is the space of all bounded $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X \times Y$. Finally if $f : X \times Y \to [0, \infty]$ is a $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function, let $f_M = M \wedge f$ so that $f_M \uparrow f$ as $M \to \infty$ and Eqs. (20.3) – (20.7) hold with f replaced by f_M for all $M \in \mathbb{N}$. Repeated use of the monotone convergence theorem allows us to pass to the limit $M \to \infty$ in these equations to deduce the theorem in the case μ and ν are finite measures. For the σ – finite case, choose $X_n \in \mathcal{M}$, $Y_n \in \mathcal{N}$ such that $X_n \uparrow X$, $Y_n \uparrow Y$, $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$ for all $m, n \in \mathbb{N}$. Then define $\mu_m(A) = \mu(X_m \cap A)$ and $\nu_n(B) = \nu(Y_n \cap B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ or equivalently $d\mu_m = 1_{X_m} d\mu$ and $d\nu_n = 1_{Y_n} d\nu$. By what we have just proved Eqs. (20.3) – (20.7) with μ replaced by μ_m and ν by ν_n for all $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable functions, $f : X \times Y \to [0, \infty]$. The validity of Eqs. (20.3) – (20.7) then follows by passing to the limits $m \to \infty$ and then $n \to \infty$ making use of the monotone convergence theorem in the form,

$$\int_X u d\mu_m = \int_X u 1_{X_m} d\mu \uparrow \int_X u d\mu \text{ as } m \to \infty$$

and

$$\int_Y v d\mu_n = \int_Y v 1_{Y_n} d\mu \uparrow \int_Y v d\mu \text{ as } n \to \infty$$

for all $u \in L^+(X, \mathcal{M})$ and $v \in L^+(Y, \mathcal{N})$. ∎

Corollary 20.6. *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ – finite measure spaces. Then there exists a unique measure π on $\mathcal{M} \otimes \mathcal{N}$ such that $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Moreover π is given by*

$$\pi(E) = \int_X d\mu(x) \int_Y d\nu(y) 1_E(x, y) = \int_Y d\nu(y) \int_X d\mu(x) 1_E(x, y) \tag{20.9}$$

for all $E \in \mathcal{M} \otimes \mathcal{N}$ and π is σ – finite.

Proof. Notice that any measure π such that $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ is necessarily σ – finite. Indeed, let $X_n \in \mathcal{M}$ and $Y_n \in \mathcal{N}$ be chosen so that $\mu(X_n) < \infty$, $\nu(Y_n) < \infty$, $X_n \uparrow X$ and $Y_n \uparrow Y$, then $X_n \times Y_n \in \mathcal{M} \otimes \mathcal{N}$, $X_n \times Y_n \uparrow X \times Y$ and $\pi(X_n \times Y_n) < \infty$ for all n. The uniqueness assertion is a consequence of Theorem 19.55 or see Theorem 25.6 below with $\mathcal{E} = \mathcal{M} \times \mathcal{N}$. For the existence, it suffices to observe, using the monotone convergence theorem, that π defined in Eq. (20.9) is a measure on $\mathcal{M} \otimes \mathcal{N}$. Moreover this measure satisfies $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ from Eq. (20.8 ∎

Notation 20.7 *The measure π is called the product measure of μ and ν and will be denoted by $\mu \otimes \nu$.*

Theorem 20.8 (Tonelli's Theorem). *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ – finite measure spaces and $\pi = \mu \otimes \nu$ is the product measure on $\mathcal{M} \otimes \mathcal{N}$. If $f \in L^+(X \times Y, \mathcal{M} \otimes \mathcal{N})$, then $f(\cdot, y) \in L^+(X, \mathcal{M})$ for all $y \in Y$, $f(x, \cdot) \in L^+(Y, \mathcal{N})$ for all $x \in X$,*

$$\int\limits_Y f(\cdot, y) d\nu(y) \in L^+(X, \mathcal{M}), \ \int\limits_X f(x, \cdot) d\mu(x) \in L^+(Y, \mathcal{N})$$

and

$$\int_{X \times Y} f \ d\pi = \int_X d\mu(x) \int_Y d\nu(y) f(x, y) \tag{20.10}$$

$$= \int_Y d\nu(y) \int_X d\mu(x) f(x, y). \tag{20.11}$$

Proof. By Theorem 20.5 and Corollary 20.6, the theorem holds when $f = 1_E$ with $E \in \mathcal{M} \otimes \mathcal{N}$. Using the linearity of all of the statements, the theorem is also true for non-negative simple functions. Then using the monotone convergence theorem repeatedly along with the approximation Theorem 18.42, one deduces the theorem for general $f \in L^+(X \times Y, \mathcal{M} \otimes \mathcal{N})$. ∎

The following convention will be in force for the rest of this chapter.

Convention: If $(X, \mathcal{M}, \mu)$ is a measure space and $f : X \to \mathbb{C}$ is a measurable but non-integrable function, i.e. $\int_X |f|\, d\mu = \infty$, by convention we will define $\int_X f d\mu := 0$. However if f is a non-negative function (i.e. $f : X \to [0, \infty]$) is a non-integrable function we will still write $\int_X f d\mu = \infty$.

Theorem 20.9 (Fubini's Theorem). *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ – finite measure spaces, $\pi = \mu \otimes \nu$ is the product measure on $\mathcal{M} \otimes \mathcal{N}$ and $f : X \times Y \to \mathbb{C}$ is a $\mathcal{M} \otimes \mathcal{N}$ – measurable function. Then the following three conditions are equivalent:*

$$\int_{X \times Y} |f|\, d\pi < \infty, \text{ i.e. } f \in L^1(\pi), \tag{20.12}$$

$$\int_X \left(\int_Y |f(x,y)|\, d\nu(y) \right) d\mu(x) < \infty \text{ and} \tag{20.13}$$

$$\int_Y \left(\int_X |f(x,y)|\, d\mu(x) \right) d\nu(y) < \infty. \tag{20.14}$$

If any one (and hence all) of these condition hold, then $f(x, \cdot) \in L^1(\nu)$ for μ a.e. x, $f(\cdot, y) \in L^1(\mu)$ for ν a.e. y, $\int_Y f(\cdot, y) dv(y) \in L^1(\mu)$, $\int_X f(x, \cdot) d\mu(x) \in L^1(\nu)$ and Eqs. (20.10) and (20.11) are still valid.

Proof. The equivalence of Eqs. (20.12) – (20.14) is a direct consequence of Tonelli's Theorem 20.8. Now suppose $f \in L^1(\pi)$ is a real valued function and let

$$E := \left\{ x \in X : \int_Y |f(x,y)|\, d\nu(y) = \infty \right\}. \tag{20.15}$$

Then by Tonelli's theorem, $x \to \int_Y |f(x,y)|\, d\nu(y)$ is measurable and hence $E \in \mathcal{M}$. Moreover Tonelli's theorem implies

$$\int_X \left[\int_Y |f(x,y)|\, d\nu(y) \right] d\mu(x) = \int_{X \times Y} |f|\, d\pi < \infty$$

which implies that $\mu(E) = 0$. Let $f_\pm$ be the positive and negative parts of f, then using the above convention we have

$$\begin{aligned} \int_Y f(x,y)\, d\nu(y) &= \int_Y 1_E(x)\, f(x,y)\, d\nu(y) \\ &= \int_Y 1_E(x) \left[f_+(x,y) - f_-(x,y) \right] d\nu(y) \\ &= \int_Y 1_E(x)\, f_+(x,y)\, d\nu(y) - \int_Y 1_E(x)\, f_-(x,y)\, d\nu(y). \end{aligned} \tag{20.16}$$

Noting that $1_E(x)\, f_\pm(x,y) = (1_E \otimes 1_Y \cdot f_\pm)(x,y)$ is a positive $\mathcal{M} \otimes \mathcal{N}$ – measurable function, it follows from another application of Tonelli's theorem that $x \to \int_Y f(x,y)\,d\nu(y)$ is $\mathcal{M}$ – measurable, being the difference of two measurable functions. Moreover

$$\int_X \left|\int_Y f(x,y)\,d\nu(y)\right| d\mu(x) \le \int_X \left[\int_Y |f(x,y)|\,d\nu(y)\right] d\mu(x) < \infty,$$

which shows $\int_Y f(\cdot,y)dv(y) \in L^1(\mu)$. Integrating Eq. (20.16) on x and using Tonelli's theorem repeatedly implies,

$$\begin{aligned}
&\int_X \left[\int_Y f(x,y)\,d\nu(y)\right] d\mu(x) \\
&\quad = \int_X d\mu(x) \int_Y d\nu(y)\, 1_E(x)\, f_+(x,y) - \int_X d\mu(x) \int_Y d\nu(y)\, 1_E(x)\, f_-(x,y) \\
&\quad = \int_Y d\nu(y) \int_X d\mu(x)\, 1_E(x)\, f_+(x,y) - \int_Y d\nu(y) \int_X d\mu(x)\, 1_E(x)\, f_-(x,y) \\
&\quad = \int_Y d\nu(y) \int_X d\mu(x)\, f_+(x,y) - \int_Y d\nu(y) \int_X d\mu(x)\, f_-(x,y) \\
&\quad = \int_{X\times Y} f_+ d\pi - \int_{X\times Y} f_- d\pi = \int_{X\times Y} (f_+ - f_-)\, d\pi = \int_{X\times Y} f d\pi \quad (20.17)
\end{aligned}$$

which proves Eq. (20.10) holds.

Now suppose that $f = u + iv$ is complex valued and again let E be as in Eq. (20.15). Just as above we still have $E \in \mathcal{M}$ and $\mu(E) < \infty$. By our convention,

$$\begin{aligned}
\int_Y f(x,y)\,d\nu(y) &= \int_Y 1_E(x)\, f(x,y)\,d\nu(y) = \int_Y 1_E(x)\,[u(x,y) + iv(x,y)]\,d\nu(y) \\
&= \int_Y 1_E(x)\, u(x,y)\,d\nu(y) + i\int_Y 1_E(x)\, v(x,y)\,d\nu(y)
\end{aligned}$$

which is measurable in x by what we have just proved. Similarly one shows $\int_Y f(\cdot,y)\,d\nu(y) \in L^1(\mu)$ and Eq. (20.10) still holds by a computation similar to that done in Eq. (20.17). The assertions pertaining to Eq. (20.11) may be proved in the same way. ■

Notation 20.10 *Given $E \subset X \times Y$ and $x \in X$, let*

$$_xE := \{y \in Y : (x,y) \in E\}.$$

Similarly if $y \in Y$ is given let

$$E_y := \{x \in X : (x,y) \in E\}.$$

If $f : X \times Y \to \mathbb{C}$ is a function let $f_x = f(x,\cdot)$ and $f^y := f(\cdot,y)$ so that $f_x : Y \to \mathbb{C}$ and $f^y : X \to \mathbb{C}$.

Theorem 20.11. *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are complete σ – finite measure spaces. Let $(X \times Y, \mathcal{L}, \lambda)$ be the completion of $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$. If f is $\mathcal{L}$ – measurable and (a) $f \geq 0$ or (b) $f \in L^1(\lambda)$ then f_x is $\mathcal{N}$ – measurable for μ a.e. x and f^y is $\mathcal{M}$ – measurable for ν a.e. y and in case (b) $f_x \in L^1(\nu)$ and $f^y \in L^1(\mu)$ for μ a.e. x and ν a.e. y respectively. Moreover,*

$$\left(x \to \int_Y f_x d\nu \right) \in L^1(\mu) \;\; \textit{and} \;\; \left(y \to \int_X f^y d\mu \right) \in L^1(\nu)$$

and

$$\int_{X \times Y} f d\lambda = \int_Y d\nu \int_X d\mu \, f = \int_X d\mu \int_Y d\nu \, f.$$

Proof. If $E \in \mathcal{M} \otimes \mathcal{N}$ is a $\mu \otimes \nu$ null set (i.e. $(\mu \otimes \nu)(E) = 0$), then

$$0 = (\mu \otimes \nu)(E) = \int\limits_X \nu({}_xE) d\mu(x) = \int\limits_X \mu(E_y) d\nu(y).$$

This shows that

$$\mu(\{x : \nu({}_xE) \neq 0\}) = 0 \text{ and } \nu(\{y : \mu(E_y) \neq 0\}) = 0,$$

i.e. $\nu({}_xE) = 0$ for μ a.e. x and $\mu(E_y) = 0$ for ν a.e. y. If h is $\mathcal{L}$ measurable and $h = 0$ for λ – a.e., then there exists $E \in \mathcal{M} \otimes \mathcal{N}$ such that $\{(x, y) : h(x, y) \neq 0\} \subset E$ and $(\mu \otimes \nu)(E) = 0$. Therefore $|h(x, y)| \leq 1_E(x, y)$ and $(\mu \otimes \nu)(E) = 0$. Since

$$\begin{aligned} \{h_x \neq 0\} &= \{y \in Y : h(x, y) \neq 0\} \subset {}_xE \text{ and} \\ \{h_y \neq 0\} &= \{x \in X : h(x, y) \neq 0\} \subset E_y \end{aligned}$$

we learn that for μ a.e. x and ν a.e. y that $\{h_x \neq 0\} \in \mathcal{M}$, $\{h_y \neq 0\} \in \mathcal{N}$, $\nu(\{h_x \neq 0\}) = 0$ and a.e. and $\mu(\{h_y \neq 0\}) = 0$. This implies $\int_Y h(x, y) d\nu(y)$ exists and equals 0 for μ a.e. x and similarly that $\int_X h(x, y) d\mu(x)$ exists and equals 0 for ν a.e. y. Therefore

$$0 = \int_{X \times Y} h d\lambda = \int_Y \left(\int_X h d\mu \right) d\nu = \int_X \left(\int_Y h d\nu \right) d\mu.$$

For general $f \in L^1(\lambda)$, we may choose $g \in L^1(\mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$ such that $f(x, y) = g(x, y)$ for $\lambda-$ a.e. (x, y). Define $h := f - g$. Then $h = 0$, $\lambda-$ a.e. Hence by what we have just proved and Theorem 20.8 $f = g + h$ has the following properties:

1. For μ a.e. x, $y \to f(x, y) = g(x, y) + h(x, y)$ is in $L^1(\nu)$ and

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

2. For ν a.e. y, $x \to f(x,y) = g(x,y) + h(x,y)$ is in $L^1(\mu)$ and

$$\int_X f(x,y)d\mu(x) = \int_X g(x,y)d\mu(x).$$

From these assertions and Theorem 20.8, it follows that

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

Similarly it is shown that

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

■

The previous theorems have obvious generalizations to products of any finite number of σ − finite measure spaces. For example the following theorem holds.

Theorem 20.12. *Suppose $\{(X_i, \mathcal{M}_i, \mu_i)\}_{i=1}^n$ are σ – finite measure spaces and $X := X_1 \times \cdots \times X_n$. Then there exists a unique measure, π, on $(X, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n)$ such that $\pi(A_1 \times \cdots \times A_n) = \mu_1(A_1) \ldots \mu_n(A_n)$ for all $A_i \in \mathcal{M}_i$. (This measure and its completion will be denote by $\mu_1 \otimes \cdots \otimes \mu_n$.) If $f : X \to [0, \infty]$ is a $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ – measurable function then*

$$\int_X f d\pi = \int_{X_{\sigma(1)}} d\mu_{\sigma(1)}(x_{\sigma(1)}) \ldots \int_{X_{\sigma(n)}} d\mu_{\sigma(n)}(x_{\sigma(n)})\ f(x_1, \ldots, x_n) \quad (20.18)$$

where σ is any permutation of $\{1, 2, \ldots, n\}$. This equation also holds for any $f \in L^1(\pi)$ and moreover, $f \in L^1(\pi)$ iff

$$\int_{X_{\sigma(1)}} d\mu_{\sigma(1)}(x_{\sigma(1)}) \ldots \int_{X_{\sigma(n)}} d\mu_{\sigma(n)}(x_{\sigma(n)})\ |f(x_1, \ldots, x_n)| < \infty$$

for some (and hence all) permutations, σ.

This theorem can be proved by the same methods as in the two factor case, see Exercise 20.4. Alternatively, one can use the theorems already proved and induction on n, see Exercise 20.5 in this regard.

Example 20.13. In this example we will show

$$\lim_{M\to\infty}\int_0^M \frac{\sin x}{x}dx = \pi/2. \tag{20.19}$$

To see this write $\frac{1}{x}=\int_0^\infty e^{-tx}dt$ and use Fubini-Tonelli to conclude that

$$\begin{aligned}\int_0^M \frac{\sin x}{x}dx &= \int_0^M\left[\int_0^\infty e^{-tx}\sin x\ dt\right]dx \\ &= \int_0^\infty\left[\int_0^M e^{-tx}\sin x\ dx\right]dt \\ &= \int_0^\infty \frac{1}{1+t^2}\left(1-te^{-Mt}\sin M - e^{-Mt}\cos M\right)dt \\ &\to \int_0^\infty \frac{1}{1+t^2}dt = \frac{\pi}{2} \text{ as } M\to\infty,\end{aligned}$$

wherein we have used the dominated convergence theorem to pass to the limit.

The next example is a refinement of this result.

Example 20.14. We have

$$\int_0^\infty \frac{\sin x}{x}e^{-\Lambda x}dx = \frac{1}{2}\pi - \arctan\Lambda \text{ for all } \Lambda>0 \tag{20.20}$$

and for$\Lambda, M\in[0,\infty)$,

$$\left|\int_0^M \frac{\sin x}{x}e^{-\Lambda x}dx - \frac{1}{2}\pi + \arctan\Lambda\right| \le C\frac{e^{-M\Lambda}}{M} \tag{20.21}$$

where $C=\max_{x\ge 0}\frac{1+x}{1+x^2}=\frac{1}{2\sqrt{2}-2}\cong 1.2$. In particular Eq. (20.19) is valid.

To verify these assertions, first notice that by the fundamental theorem of calculus,

$$|\sin x| = \left|\int_0^x \cos y dy\right| \le \left|\int_0^x |\cos y|\, dy\right| \le \left|\int_0^x 1dy\right| = |x|$$

so $\left|\frac{\sin x}{x}\right|\le 1$ for all $x\neq 0$. Making use of the identity

$$\int_0^\infty e^{-tx}dt = 1/x$$

and Fubini's theorem,

$$\begin{aligned}
\int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx &= \int_0^M dx \sin x\, e^{-\Lambda x} \int_0^\infty e^{-tx} dt \\
&= \int_0^\infty dt \int_0^M dx \sin x\, e^{-(\Lambda+t)x} \\
&= \int_0^\infty \frac{1 - (\cos M + (\Lambda+t)\sin M)\, e^{-M(\Lambda+t)}}{(\Lambda+t)^2 + 1} dt \\
&= \int_0^\infty \frac{1}{(\Lambda+t)^2+1} dt - \int_0^\infty \frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1} e^{-M(\Lambda+t)} dt \\
&= \frac{1}{2}\pi - \arctan \Lambda - \varepsilon(M,\Lambda) \qquad (20.22)
\end{aligned}$$

where

$$\varepsilon(M,\Lambda) = \int_0^\infty \frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1} e^{-M(\Lambda+t)} dt.$$

Since

$$\left|\frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1}\right| \le \frac{1+(\Lambda+t)}{(\Lambda+t)^2+1} \le C,$$

$$|\varepsilon(M,\Lambda)| \le \int_0^\infty e^{-M(\Lambda+t)} dt = C\frac{e^{-M\Lambda}}{M}.$$

This estimate along with Eq. (20.22) proves Eq. (20.21) from which Eq. (20.19) follows by taking $\Lambda \to \infty$ and Eq. (20.20) follows (using the dominated convergence theorem again) by letting $M \to \infty$.

20.2 Lebesgue Measure on $\mathbb{R}^d$ and the Change of Variables Theorem

Notation 20.15 *Let*

$$m^d := \overbrace{m \otimes \cdots \otimes m}^{d\ times} \text{ on } \mathcal{B}_{\mathbb{R}^d} = \overbrace{\mathcal{B}_{\mathbb{R}} \otimes \cdots \otimes \mathcal{B}_{\mathbb{R}}}^{d\ times}$$

be the d – fold product of Lebesgue measure m on $\mathcal{B}_{\mathbb{R}}$. We will also use m^d to denote its completion and let $\mathcal{L}_d$ be the completion of $\mathcal{B}_{\mathbb{R}^d}$ relative to m^d. A subset $A \in \mathcal{L}_d$ is called a Lebesgue measurable set and m^d is called d – dimensional Lebesgue measure, or just Lebesgue measure for short.

Definition 20.16. *A function $f : \mathbb{R}^d \to \mathbb{R}$ is* ***Lebesgue measurable*** *if $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{L}_d$.*

Notation 20.17 *I will often be sloppy in the sequel and write m for m^d and dx for $dm(x) = dm^d(x)$, i.e.*

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_{\mathbb{R}^d} f dm = \int_{\mathbb{R}^d} f dm^d.$$

Hopefully the reader will understand the meaning from the context.

Theorem 20.18. *Lebesgue measure m^d is translation invariant. Moreover m^d is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0,1]^d) = 1$.*

Proof. Let $A = J_1 \times \cdots \times J_d$ with $J_i \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}^d$. Then

$$x + A = (x_1 + J_1) \times (x_2 + J_2) \times \cdots \times (x_d + J_d)$$

and therefore by translation invariance of m on $\mathcal{B}_{\mathbb{R}}$ we find that

$$m^d(x + A) = m(x_1 + J_1) \dots m(x_d + J_d) = m(J_1) \dots m(J_d) = m^d(A)$$

and hence $m^d(x + A) = m^d(A)$ for all $A \in \mathcal{B}_{\mathbb{R}^d}$ by Corollary 19.57. From this fact we see that the measure $m^d(x + \cdot)$ and $m^d(\cdot)$ have the same null sets. Using this it is easily seen that $m(x + A) = m(A)$ for all $A \in \mathcal{L}_d$. The proof of the second assertion is Exercise 20.6. ∎

Theorem 20.19 (Change of Variables Theorem). *Let $\Omega \subset_o \mathbb{R}^d$ be an open set and $T : \Omega \to T(\Omega) \subset_o \mathbb{R}^d$ be a C^1 – diffeomorphism,*[1] *see Figure 20.1. Then for any Borel measurable function, $f : T(\Omega) \to [0, \infty]$,*

$$\int_{\Omega} f(T(x))\,|\det T'(x)|\,dx = \int_{T(\Omega)} f(y)\,dy, \tag{20.23}$$

where $T'(x)$ is the linear transformation on $\mathbb{R}^d$ defined by $T'(x)v := \frac{d}{dt}|_0 T(x + tv)$. More explicitly, viewing vectors in $\mathbb{R}^d$ as columns, $T'(x)$ may be represented by the matrix

$$T'(x) = \begin{bmatrix} \partial_1 T_1(x) & \dots & \partial_d T_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 T_d(x) & \dots & \partial_d T_d(x) \end{bmatrix}, \tag{20.24}$$

i.e. the i - j – matrix entry of $T'(x)$ is given by $T'(x)_{ij} = \partial_i T_j(x)$ where $T(x) = (T_1(x), \dots, T_d(x))^{\mathrm{tr}}$ and $\partial_i = \partial/\partial x_i$.

Remark 20.20. Theorem 20.19 is best remembered as the statement: if we make the change of variables $y = T(x)$, then $dy = |\det T'(x)|\,dx$. As usual, you must also change the limits of integration appropriately, i.e. if x ranges through Ω then y must range through $T(\Omega)$.

[1] That is $T : \Omega \to T(\Omega) \subset_o \mathbb{R}^d$ is a continuously differentiable bijection and the inverse map $T^{-1} : T(\Omega) \to \Omega$ is also continuously differentiable.

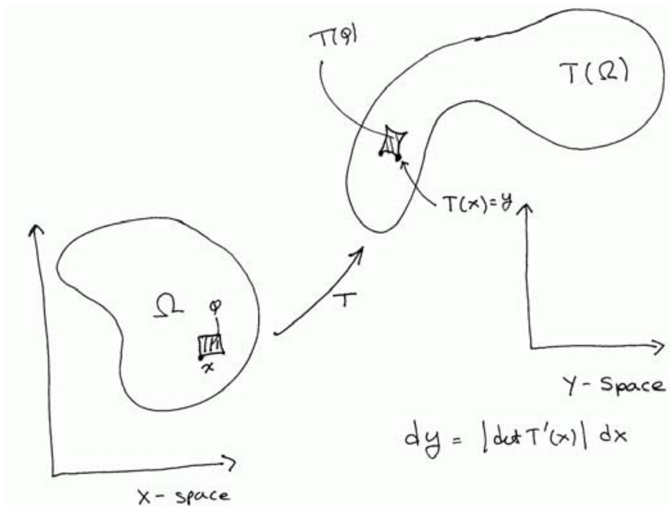


Fig. 20.1. The geometric setup of Theorem 20.19.

Proof. The proof will be by induction on d. The case $d = 1$ was essentially done in Exercise 19.8. Nevertheless, for the sake of completeness let us give a proof here. Suppose $d = 1$, $a < \alpha < \beta < b$ such that $[a, b]$ is a compact subinterval of Ω. Then $|\det T'| = |T'|$ and

$$\int_{[a,b]} 1_{T((\alpha,\beta])}(T(x))\,|T'(x)|\,dx = \int_{[a,b]} 1_{(\alpha,\beta]}(x)\,|T'(x)|\,dx = \int_\alpha^\beta |T'(x)|\,dx.$$

If $T'(x) > 0$ on $[a, b]$, then

$$\begin{aligned}\int_\alpha^\beta |T'(x)|\,dx &= \int_\alpha^\beta T'(x)\,dx = T(\beta) - T(\alpha)\\ &= m(T((\alpha,\beta])) = \int_{T([a,b])} 1_{T((\alpha,\beta])}(y)\,dy\end{aligned}$$

while if $T'(x) < 0$ on $[a, b]$, then

$$\begin{aligned}\int_\alpha^\beta |T'(x)|\,dx &= -\int_\alpha^\beta T'(x)\,dx = T(\alpha) - T(\beta)\\ &= m(T((\alpha,\beta])) = \int_{T([a,b])} 1_{T((\alpha,\beta])}(y)\,dy.\end{aligned}$$

Combining the previous three equations shows

$$\int_{[a,b]} f(T(x))\,|T'(x)|\,dx = \int_{T([a,b])} f(y)\,dy \tag{20.25}$$

whenever f is of the form $f = 1_{T((\alpha,\beta])}$ with $a < \alpha < \beta < b$. An application of Dynkin's multiplicative system Theorem 18.51 then implies that Eq. (20.25) holds for every bounded measurable function $f : T([a,b]) \to \mathbb{R}$. (Observe that $|T'(x)|$ is continuous and hence bounded for x in the compact interval, $[a,b]$. From Exercise 10.12, $\Omega = \coprod_{n=1}^N (a_n, b_n)$ where $a_n, b_n \in \mathbb{R}\cup\{\pm\infty\}$ for $n = 1, 2, \dots < N$ with $N = \infty$ possible. Hence if $f : T(\Omega) \to \mathbb{R}_+$ is a Borel measurable function and $a_n < \alpha_k < \beta_k < b_n$ with $\alpha_k \downarrow a_n$ and $\beta_k \uparrow b_n$, then by what we have already proved and the monotone convergence theorem

$$\begin{aligned}
\int_\Omega 1_{(a_n,b_n)} \cdot (f\circ T) \cdot |T'| dm &= \int_\Omega \left(1_{T((a_n,b_n))} \cdot f\right) \circ T \cdot |T'| dm \\
&= \lim_{k\to\infty} \int_\Omega \left(1_{T([\alpha_k,\beta_k])} \cdot f\right) \circ T \cdot |T'|\ dm \\
&= \lim_{k\to\infty} \int_{T(\Omega)} 1_{T([\alpha_k,\beta_k])} \cdot f\ dm \\
&= \int_{T(\Omega)} 1_{T((a_n,b_n))} \cdot f\ dm.
\end{aligned}$$

Summing this equality on n, then shows Eq. (20.23) holds.

To carry out the induction step, we now suppose $d > 1$ and suppose the theorem is valid with d being replaced by $d-1$. For notational compactness, let us write vectors in $\mathbb{R}^d$ as row vectors rather than column vectors. Nevertheless, the matrix associated to the differential, $T'(x)$, will always be taken to be given as in Eq. (20.24).

Case 1. Suppose $T(x)$ has the form

$$T(x) = (x_i, T_2(x), \dots, T_d(x)) \tag{20.26}$$

or

$$T(x) = (T_1(x), \dots, T_{d-1}(x), x_i) \tag{20.27}$$

for some $i \in \{1, \dots, d\}$. For definiteness we will assume T is as in Eq. (20.26), the case of T in Eq. (20.27) may be handled similarly. For $t \in \mathbb{R}$, let $i_t : \mathbb{R}^{d-1} \to \mathbb{R}^d$ be the inclusion map defined by

$$i_t(w) := w_t := (w_1, \dots, w_{i-1}, t, w_{i+1}, \dots, w_{d-1}),$$

Ω_t be the (possibly empty) open subset of $\mathbb{R}^{d-1}$ defined by

$$\Omega_t := \left\{w \in \mathbb{R}^{d-1} : (w_1, \dots, w_{i-1}, t, w_{i+1}, \dots, w_{d-1}) \in \Omega\right\}$$

and $T_t : \Omega_t \to \mathbb{R}^{d-1}$ be defined by

$$T_t(w) = (T_2(w_t), \dots, T_d(w_t)),$$

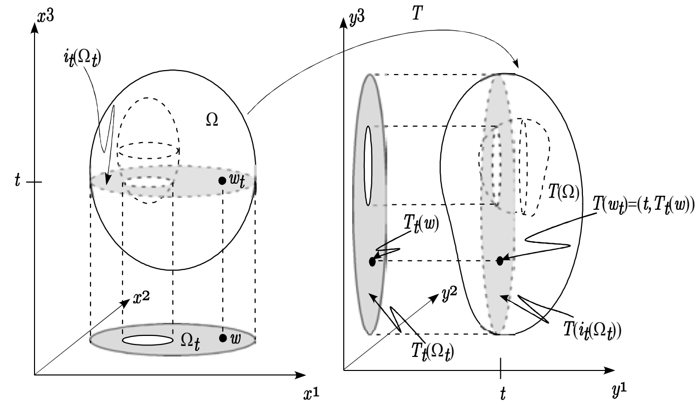


Fig. 20.2. In this picture $d = i = 3$ and Ω is an egg-shaped region with an egg-shaped hole. The picture indicates the geometry associated with the map T and slicing the set Ω along planes where $x_3 = t$.

see Figure 20.2. Expanding $\det T'(w_t)$ along the first row of the matrix $T'(w_t)$ shows

$$|\det T'(w_t)| = |\det T'_t(w)|.$$

Now by the Fubini-Tonelli Theorem and the induction hypothesis,

$$\begin{aligned}
\int_\Omega f\circ T|\det T'|dm &= \int_{\mathbb{R}^d} 1_\Omega \cdot f\circ T|\det T'|dm \\
&= \int_{\mathbb{R}^d} 1_\Omega(w_t)(f\circ T)(w_t)|\det T'(w_t)|dwdt \\
&= \int_{\mathbb{R}}\left[\int_{\Omega_t}(f\circ T)(w_t)|\det T'(w_t)|dw\right]dt \\
&= \int_{\mathbb{R}}\left[\int_{\Omega_t} f(t,T_t(w))|\det T'_t(w)|dw\right]dt \\
&= \int_{\mathbb{R}}\left[\int_{T_t(\Omega_t)} f(t,z)\,dz\right]dt = \int_{\mathbb{R}}\left[\int_{\mathbb{R}^{d-1}} 1_{T(\Omega)}(t,z)\,f(t,z)\,dz\right]dt \\
&= \int_{T(\Omega)} f(y)\,dy
\end{aligned}$$

wherein the last two equalities we have used Fubini-Tonelli along with the identity;

$$T(\Omega) = \coprod_{t \in \mathbb{R}} T(i_t(\Omega)) = \coprod_{t \in \mathbb{R}} \{(t, z) : z \in T_t(\Omega_t)\}.$$

Case 2. (Eq. (20.23) is true locally.) Suppose that $T : \Omega \to \mathbb{R}^d$ is a general map as in the statement of the theorem and $x_0 \in \Omega$ is an arbitrary point. We will now show there exists an open neighborhood $W \subset \Omega$ of x_0 such that

$$\int_W f \circ T |\det T'| dm = \int_{T(W)} f dm$$

holds for all Borel measurable function, $f : T(W) \to [0, \infty]$. Let M_i be the 1-i minor of $T'(x_0)$, i.e. the determinant of $T'(x_0)$ with the first row and i^{th} – column removed. Since

$$0 \neq \det T'(x_0) = \sum_{i=1}^{d} (-1)^{i+1} \partial_i T_j(x_0) \cdot M_i,$$

there must be some i such that $M_i \neq 0$. Fix an i such that $M_i \neq 0$ and let,

$$S(x) := (x_i, T_2(x), \ldots, T_d(x)). \tag{20.28}$$

Observe that $|\det S'(x_0)| = |M_i| \neq 0$. Hence by the inverse function Theorem 16.25, there exist an open neighborhood W of x_0 such that $W \subset_o \Omega$ and $S(W) \subset_o \mathbb{R}^d$ and $S : W \to S(W)$ is a C^1 – diffeomorphism. Let $R : S(W) \to T(W) \subset_o \mathbb{R}^d$ to be the C^1 – diffeomorphism defined by

$$R(z) := T \circ S^{-1}(z) \text{ for all } z \in S(W).$$

Because

$$(T_1(x), \ldots, T_d(x)) = T(x) = R(S(x)) = R((x_i, T_2(x), \ldots, T_d(x)))$$

for all $x \in W$, if

$$(z_1, z_2, \ldots, z_d) = S(x) = (x_i, T_2(x), \ldots, T_d(x))$$

then

$$R(z) = \left(T_1\left(S^{-1}(z)\right), z_2, \ldots, z_d\right). \tag{20.29}$$

Observe that S is a map of the form in Eq. (20.26), R is a map of the form in Eq. (20.27), $T'(x) = R'(S(x)) S'(x)$ (by the chain rule) and (by the multiplicative property of the determinant)

$$|\det T'(x)| = |\det R'(S(x))| \, |\det S'(x)| \;\; \forall \, x \in W.$$

So if $f : T(W) \to [0, \infty]$ is a Borel measurable function, two applications of the results in Case 1. shows,

$$\begin{aligned}\int_W f\circ T\cdot|\det T'|dm &= \int_W (f\circ R\cdot|\det R'|)\circ S\cdot|\det S'|\ \ dm \\ &= \int_{S(W)} f\circ R\cdot|\det R'|dm = \int_{R(S(W))} fdm \\ &= \int_{T(W)} fdm\end{aligned}$$

and Case 2. is proved.

Case 3. (General Case.) Let $f:\Omega\to[0,\infty]$ be a general non-negative Borel measurable function and let

$$K_n := \{x\in\Omega : \operatorname{dist}(x,\Omega^c)\geq 1/n \text{ and } |x|\leq n\}.$$

Then each K_n is a compact subset of Ω and $K_n\uparrow\Omega$ as $n\to\infty$. Using the compactness of K_n and case 2, for each $n\in\mathbb{N}$, there is a finite open cover $\mathcal{W}_n$ of K_n such that $W\subset\Omega$ and Eq. (20.23) holds with Ω replaced by W for each $W\in\mathcal{W}_n$. Let $\{W_i\}_{i=1}^\infty$ be an enumeration of $\cup_{n=1}^\infty\mathcal{W}_n$ and set $\tilde{W}_1 = W_1$ and $\tilde{W}_i := W_i\setminus(W_1\cup\dots\cup W_{i-1})$ for all $i\geq 2$. Then $\Omega = \coprod_{i=1}^\infty\tilde{W}_i$ and by repeated use of case 2.,

$$\begin{aligned}\int_\Omega f\circ T|\det T'|dm &= \sum_{i=1}^\infty\int_\Omega 1_{\tilde{W}_i}\cdot(f\circ T)\cdot|\det T'|dm \\ &= \sum_{i=1}^\infty\int_{W_i}\left[\left(1_{T(\tilde{W}_i)}f\right)\circ T\right]\cdot|\det T'|dm \\ &= \sum_{i=1}^\infty\int_{T(W_i)} 1_{T(\tilde{W}_i)}\cdot f\ dm = \sum_{i=1}^n\int_{T(\Omega)} 1_{T(\tilde{W}_i)}\cdot f\ dm \\ &= \int_{T(\Omega)} fdm.\end{aligned}$$

■

Remark 20.21. When $d=1$, one often learns the change of variables formula as

$$\int_a^b f(T(x))\,T'(x)\,dx = \int_{T(a)}^{T(b)} f(y)\,dy \tag{20.30}$$

where $f:[a,b]\to\mathbb{R}$ is a continuous function and T is C^1 – function defined in a neighborhood of $[a,b]$. If $T'>0$ on (a,b) then $T((a,b)) = (T(a),T(b))$ and Eq. (20.30) is implies Eq. (20.23) with $\Omega=(a,b)$. On the other hand if $T'<0$ on (a,b) then $T((a,b)) = (T(b),T(a))$ and Eq. (20.30) is equivalent to

$$\int_{(a,b)} f\left(T\left(x\right)\right)\left(-\left|T'\left(x\right)\right|\right)dx = -\int_{T(b)}^{T(a)} f\left(y\right)dy = -\int_{T((a,b))} f\left(y\right)dy$$

which is again implies Eq. (20.23). On the other hand Eq. Eq. (20.30) is more general than Eq. (20.23) since it does not require T to be injective. The standard proof of Eq. (20.30) is as follows. For $z \in T\left([a,b]\right)$, let

$$F\left(z\right) := \int_{T(a)}^{z} f\left(y\right)dy.$$

Then by the chain rule and the fundamental theorem of calculus,

$$\begin{aligned}\int_a^b f\left(T\left(x\right)\right)T'\left(x\right)dx &= \int_a^b F'\left(T\left(x\right)\right)T'\left(x\right)dx = \int_a^b \frac{d}{dx}\left[F\left(T\left(x\right)\right)\right]dx \\ &= F\left(T\left(x\right)\right)|_a^b = \int_{T(a)}^{T(b)} f\left(y\right)dy.\end{aligned}$$

An application of Dynkin's multiplicative systems theorem (in the form of Corollary 18.55) now shows that Eq. (20.30) holds for all bounded measurable functions f on (a,b). Then by the usual truncation argument, it also holds for all positive measurable functions on (a,b).

Example 20.22. Continuing the setup in Theorem 20.19, if $A \in \mathcal{B}_\Omega$, then

$$\begin{aligned}m\left(T\left(A\right)\right) &= \int_{\mathbb{R}^d} 1_{T(A)}\left(y\right)dy = \int_{\mathbb{R}^d} 1_{T(A)}\left(Tx\right)\left|\det T'\left(x\right)\right|dx \\ &= \int_{\mathbb{R}^d} 1_A\left(x\right)\left|\det T'\left(x\right)\right|dx\end{aligned}$$

wherein the second equality we have made the change of variables, $y = T\left(x\right)$. Hence we have shown

$$d\left(m \circ T\right) = \left|\det T'\left(\cdot\right)\right|dm.$$

In particular if $T \in GL(d,\mathbb{R}) = GL(\mathbb{R}^d)$ – the space of $d \times d$ invertible matrices, then $m \circ T = \left|\det T\right| m$, i.e.

$$m\left(T\left(A\right)\right) = \left|\det T\right| m\left(A\right) \text{ for all} A \in \mathcal{B}_{\mathbb{R}^d}. \tag{20.31}$$

This equation also shows that $m \circ T$ and m have the same null sets and hence the equality in Eq. (20.31) is valid for any $A \in \mathcal{L}_d$.

Exercise 20.1. Show that $f \in L^1\left(T\left(\Omega\right), m^d\right)$ iff

$$\int_\Omega \left|f \circ T\right| \left|\det T'\right| dm < \infty$$

and if $f \in L^1\left(T\left(\Omega\right), m^d\right)$, then Eq. (20.23) holds.

Example 20.23 (Polar Coordinates). Suppose $T : (0,\infty)\times(0,2\pi)\to\mathbb{R}^2$ is defined by

$$x = T(r,\theta) = (r\cos\theta, r\sin\theta),$$

i.e. we are making the change of variable,

$$x_1 = r\cos\theta \text{ and } x_2 = r\sin\theta \text{ for } 0<r<\infty \text{ and } 0<\theta<2\pi.$$

In this case

$$T'(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

and therefore

$$dx = |\det T'(r,\theta)|\, drd\theta = rdrd\theta.$$

Observing that

$$\mathbb{R}^2 \setminus T((0,\infty)\times(0,2\pi)) = \ell := \{(x,0) : x \geq 0\}$$

has m^2 – measure zero, it follows from the change of variables Theorem 20.19 that

$$\int_{\mathbb{R}^2} f(x)dx = \int_0^{2\pi} d\theta \int_0^\infty dr\ r\cdot f(r(\cos\theta,\sin\theta)) \tag{20.32}$$

for any Borel measurable function $f:\mathbb{R}^2\to[0,\infty]$.

Example 20.24 (Holomorphic Change of Variables). Suppose that $f : \Omega \subset_o \mathbb{C}\cong\mathbb{R}^2 \to \mathbb{C}$ is an injective holomorphic function such that $f'(z)\neq 0$ for all $z\in\Omega$. We may express f as

$$f(x+iy) = U(x,y) + iV(x,y)$$

for all $z = x+iy\in\Omega$. Hence if we make the change of variables,

$$w = u+iv = f(x+iy) = U(x,y)+iV(x,y)$$

then

$$dudv = \left|\det\begin{bmatrix} U_x & U_y \\ V_x & V_y\end{bmatrix}\right| dxdy = |U_xV_y - U_yV_x|\, dxdy.$$

Recalling that U and V satisfy the Cauchy Riemann equations, $U_x = V_y$ and $U_y = -V_x$ with $f' = U_x + iV_x$, we learn

$$U_xV_y - U_yV_x = U_x^2 + V_x^2 = |f'|^2.$$

Therefore

$$dudv = |f'(x+iy)|^2\, dxdy.$$

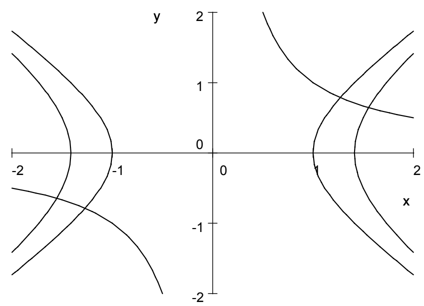


Fig. 20.3. The region Ω consists of the two curved rectangular regions shown.

Example 20.25. In this example we will evaluate the integral

$$I := \iint_{\Omega} \left(x^4 - y^4\right) dxdy$$

where

$$\Omega = \left\{(x,y) : 1 < x^2 - y^2 < 2,\ 0 < xy < 1\right\},$$

see Figure 20.3 We are going to do this by making the change of variables,

$$(u,v) := T(x,y) = \left(x^2 - y^2, xy\right),$$

in which case

$$dudv = \left|\det \begin{bmatrix} 2x & -2y \\ y & x \end{bmatrix}\right| dxdy = 2\left(x^2 + y^2\right) dxdy$$

Notice that

$$\left(x^4 - y^4\right) = \left(x^2 - y^2\right)\left(x^2 + y^2\right) = u\left(x^2 + y^2\right) = \frac{1}{2} ududv.$$

The function T is not injective on Ω but it is injective on each of its connected components. Let D be the connected component in the first quadrant so that $\Omega = -D \cup D$ and $T(\pm D) = (1,2) \times (0,1)$ The change of variables theorem then implies

$$I_{\pm} := \iint_{\pm D} \left(x^4 - y^4\right) dxdy = \frac{1}{2} \iint_{(1,2)\times(0,1)} ududv = \frac{1}{2} \frac{u^2}{2}\Big|_1^2 \cdot 1 = \frac{3}{4}$$

and therefore $I = I_+ + I_- = 2 \cdot (3/4) = 3/2$.

Exercise 20.2 (Spherical Coordinates). Let $T : (0,\infty)\times(0,\pi)\times(0,2\pi) \to \mathbb{R}^3$ be defined by

$$\begin{aligned} T\left(r,\phi,\theta\right) &= \left(r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi\right) \\ &= r\left(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi\right), \end{aligned}$$

see Figure 20.4. By making the change of variables $x = T\left(r,\phi,\theta\right)$, show

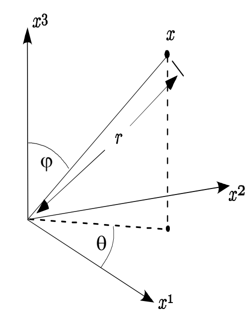


Fig. 20.4. The relation of x to (r,ϕ,θ) in spherical coordinates.

$$\int_{\mathbb{R}^3} f(x)dx = \int_0^{\pi} d\phi \int_0^{2\pi} d\theta \int_0^{\infty} dr\ r^2\sin\phi \cdot f(T\left(r,\phi,\theta\right))$$

for any Borel measurable function, $f : \mathbb{R}^3 \to [0,\infty]$.

Lemma 20.26. *Let $a > 0$ and*

$$I_d(a) := \int_{\mathbb{R}^d} e^{-a|x|^2} dm(x).$$

Then $I_d(a) = (\pi/a)^{d/2}$.

Proof. By Tonelli's theorem and induction,

$$\begin{aligned} I_d(a) &= \int_{\mathbb{R}^{d-1}\times\mathbb{R}} e^{-a|y|^2} e^{-at^2} m_{d-1}(dy)\ dt \\ &= I_{d-1}(a) I_1(a) = I_1^d(a). \end{aligned} \tag{20.33}$$

So it suffices to compute:

$$I_2(a) = \int_{\mathbb{R}^2} e^{-a|x|^2} dm(x) = \int_{\mathbb{R}^2\setminus\{0\}} e^{-a(x_1^2+x_2^2)} dx_1 dx_2.$$

Using polar coordinates, see Eq. (20.32), we find,

$$\begin{aligned} I_2(a) &= \int_0^\infty dr\ r \int_0^{2\pi} d\theta\ e^{-ar^2} = 2\pi \int_0^\infty re^{-ar^2} dr \\ &= 2\pi \lim_{M\to\infty} \int_0^M re^{-ar^2} dr = 2\pi \lim_{M\to\infty} \frac{e^{-ar^2}}{-2a}\int_0^M = \frac{2\pi}{2a} = \pi/a. \end{aligned}$$

This shows that $I_2(a) = \pi/a$ and the result now follows from Eq. (20.33). ■

20.3 The Polar Decomposition of Lebesgue Measure

Let

$$S^{d-1} = \{x \in \mathbb{R}^d : |x|^2 := \sum_{i=1}^d x_i^2 = 1\}$$

be the unit sphere in $\mathbb{R}^d$ equipped with its Borel σ – algebra, $\mathcal{B}_{S^{d-1}}$ and $\Phi : \mathbb{R}^d \setminus \{0\} \to (0,\infty) \times S^{d-1}$ be defined by $\Phi(x) := (|x|, |x|^{-1} x)$. The inverse map, $\Phi^{-1} : (0,\infty) \times S^{d-1} \to \mathbb{R}^d \setminus \{0\}$, is given by $\Phi^{-1}(r,\omega) = r\omega$. Since Φ and Φ^{-1} are continuous, they are both Borel measurable. For $E \in \mathcal{B}_{S^{d-1}}$ and $a > 0$, let

$$E_a := \{r\omega : r \in (0,a] \text{ and } \omega \in E\} = \Phi^{-1}((0,a] \times E) \in \mathcal{B}_{\mathbb{R}^d}.$$

Definition 20.27. *For $E \in \mathcal{B}_{S^{d-1}}$, let $\sigma(E) := d \cdot m(E_1)$. We call σ the surface measure on S^{d-1}.*

It is easy to check that σ is a measure. Indeed if $E \in \mathcal{B}_{S^{d-1}}$, then $E_1 = \Phi^{-1}((0,1] \times E) \in \mathcal{B}_{\mathbb{R}^d}$ so that $m(E_1)$ is well defined. Moreover if $E = \coprod_{i=1}^\infty E_i$, then $E_1 = \coprod_{i=1}^\infty (E_i)_1$ and

$$\sigma(E) = d \cdot m(E_1) = \sum_{i=1}^\infty m((E_i)_1) = \sum_{i=1}^\infty \sigma(E_i).$$

The intuition behind this definition is as follows. If $E \subset S^{d-1}$ is a set and $\varepsilon > 0$ is a small number, then the volume of

$$(1, 1+\varepsilon] \cdot E = \{r\omega : r \in (1, 1+\varepsilon] \text{ and } \omega \in E\}$$

should be approximately given by $m((1, 1+\varepsilon] \cdot E) \cong \sigma(E)\varepsilon$, see Figure 20.5 below. On the other hand

$$m((1, 1+\varepsilon]E) = m(E_{1+\varepsilon} \setminus E_1) = \left\{(1+\varepsilon)^d - 1\right\} m(E_1).$$

Therefore we expect the area of E should be given by

$$\sigma(E) = \lim_{\varepsilon\downarrow 0} \frac{\left\{(1+\varepsilon)^d - 1\right\} m(E_1)}{\varepsilon} = d \cdot m(E_1).$$

The following theorem is motivated by Example 20.23 and Exercise 20.2.

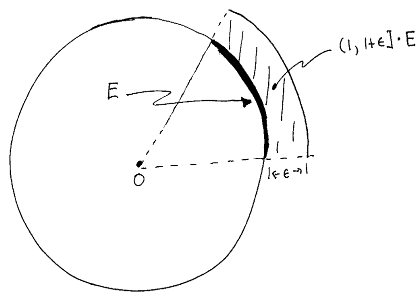


Fig. 20.5. Motivating the definition of surface measure for a sphere.

Theorem 20.28 (Polar Coordinates). *If $f : \mathbb{R}^d \to [0,\infty]$ is a $(\mathcal{B}_{R^d}, \mathcal{B})$–measurable function then*

$$\int\limits_{\mathbb{R}^d} f(x)dm(x) = \int\limits_{(0,\infty)\times S^{d-1}} f(r\omega)r^{d-1}\ drd\sigma(\omega). \tag{20.34}$$

Proof. By Exercise 19.7,

$$\int\limits_{\mathbb{R}^d} fdm = \int\limits_{\mathbb{R}^d\setminus\{0\}} \left(f\circ\Phi^{-1}\right)\circ\Phi\ dm = \int\limits_{(0,\infty)\times S^{d-1}} \left(f\circ\Phi^{-1}\right)\ d\left(\Phi_* m\right) \tag{20.35}$$

and therefore to prove Eq. (20.34) we must work out the measure $\Phi_* m$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ defined by

$$\Phi_* m(A) := m\left(\Phi^{-1}(A)\right)\ \forall\ A \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}. \tag{20.36}$$

If $A = (a,b] \times E$ with $0 < a < b$ and $E \in \mathcal{B}_{S^{d-1}}$, then

$$\Phi^{-1}(A) = \{r\omega : r \in (a,b] \text{ and } \omega \in E\} = bE_1 \setminus aE_1$$

wherein we have used $E_a = aE_1$ in the last equality. Therefore by the basic scaling properties of m and the fundamental theorem of calculus,

$$\begin{aligned}(\Phi_* m)\left((a,b]\times E\right) &= m\left(bE_1 \setminus aE_1\right) = m(bE_1) - m(aE_1) \\ &= b^d m(E_1) - a^d m(E_1) = d\cdot m(E_1)\int_a^b r^{d-1}dr.\end{aligned} \tag{20.37}$$

Letting $d\rho(r) = r^{d-1}dr$, i.e.

$$\rho(J) = \int_J r^{d-1}dr\ \forall\ J \in \mathcal{B}_{(0,\infty)}, \tag{20.38}$$

Eq. (20.37) may be written as

$$(\Phi_* m)((a,b] \times E) = \rho((a,b]) \cdot \sigma(E) = (\rho \otimes \sigma)((a,b] \times E). \tag{20.39}$$

Since

$$\mathcal{E} = \left\{(a,b] \times E : 0 < a < b \text{ and } E \in \mathcal{B}_{S^{d-1}}\right\},$$

is a π class (in fact it is an elementary class) such that $\sigma(\mathcal{E}) = \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$, it follows from Theorem 19.55 and Eq. (20.39) that $\Phi_* m = \rho \otimes \sigma$. Using this result in Eq. (20.35) gives

$$\int_{\mathbb{R}^d} f dm = \int_{(0,\infty)\times S^{d-1}} \left(f \circ \Phi^{-1}\right) \, d(\rho \otimes \sigma)$$

which combined with Tonelli's Theorem 20.8 proves Eq. (20.35). ∎

Corollary 20.29. *The surface area $\sigma(S^{d-1})$ of the unit sphere $S^{d-1} \subset \mathbb{R}^d$ is*

$$\sigma(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{20.40}$$

where Γ is the gamma function given by

$$\Gamma(x) := \int_0^\infty u^{x-1} e^{-u} dr \tag{20.41}$$

Moreover, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$.

Proof. Using Theorem 20.28 we find

$$I_d(1) = \int_0^\infty dr \; r^{d-1} e^{-r^2} \int_{S^{d-1}} d\sigma = \sigma(S^{d-1}) \int_0^\infty r^{d-1} e^{-r^2} dr.$$

We simplify this last integral by making the change of variables $u = r^2$ so that $r = u^{1/2}$ and $dr = \frac{1}{2} u^{-1/2} du$. The result is

$$\begin{aligned}\int_0^\infty r^{d-1} e^{-r^2} dr &= \int_0^\infty u^{\frac{d-1}{2}} e^{-u} \frac{1}{2} u^{-1/2} du \\ &= \frac{1}{2} \int_0^\infty u^{\frac{d}{2}-1} e^{-u} du = \frac{1}{2}\Gamma(d/2). \end{aligned} \tag{20.42}$$

Combing the the last two equations with Lemma 20.26 which states that $I_d(1) = \pi^{d/2}$, we conclude that

$$\pi^{d/2} = I_d(1) = \frac{1}{2}\sigma(S^{d-1})\Gamma(d/2)$$

which proves Eq. (20.40). Example 19.24 implies $\Gamma(1) = 1$ and from Eq. (20.42),

$$\begin{aligned}\Gamma(1/2) &= 2\int_0^\infty e^{-r^2}dr = \int_{-\infty}^\infty e^{-r^2}dr \\ &= I_1(1) = \sqrt{\pi}.\end{aligned}$$

The relation, $\Gamma(x+1) = x\Gamma(x)$ is the consequence of the following integration by parts argument:

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty e^{-u}\ u^{x+1}\ \frac{du}{u} = \int_0^\infty u^x \left(-\frac{d}{du}\ e^{-u}\right) du \\ &= x\int_0^\infty u^{x-1}\ e^{-u}\ du = x\ \Gamma(x).\end{aligned}$$

■

BRUCE: add Morrey's Inequality 72.1 here.

20.4 More proofs of the classical Weierstrass approximation Theorem 8.34

In each of these proofs we will use the reduction explained the previous proof of Theorem 8.34 to reduce to the case where $f \in C([0,1]^d)$. The first proof we will give here is based on the "weak law" of large numbers. The second will be another approximate δ – function argument.

Proof. of Theorem 8.34. Let $\mathbf{0} : = (0,0,\dots,0)$, $\mathbf{1} : = (1,1,\dots,1)$ and $[\mathbf{0},\mathbf{1}] := [0,1]^d$. By considering the real and imaginary parts of f separately, it suffices to assume $f \in C([\mathbf{0},\mathbf{1}],\mathbb{R})$. For $x \in [0,1]$, let ν_x be the measure on $\{0,1\}$ such that $\nu_x(\{0\}) = 1-x$ and $\nu_x(\{1\}) = x$. Then

$$\int_{\{0,1\}} y d\nu_x(y) = 0\cdot(1-x) + 1\cdot x = x \text{ and} \tag{20.43}$$

$$\int_{\{0,1\}} (y-x)^2 d\nu_x(y) = x^2(1-x) + (1-x)^2\cdot x = x(1-x). \tag{20.44}$$

For $x \in [\mathbf{0},\mathbf{1}]$ let $\mu_x = \nu_{x_1}\otimes\cdots\otimes\nu_{x_d}$ be the product of $\nu_{x_1},\dots,\nu_{x_d}$ on $\Omega := \{0,1\}^d$. Alternatively the measure μ_x may be described by

$$\mu_x(\{\varepsilon\}) = \prod_{i=1}^d (1-x_i)^{1-\varepsilon_i}\, x_i^{\varepsilon_i} \tag{20.45}$$

for $\varepsilon \in \Omega$. Notice that $\mu_x(\{\varepsilon\})$ is a degree d polynomial in x for each $\varepsilon \in \Omega$. For $n \in \mathbb{N}$ and $x \in [\mathbf{0},\mathbf{1}]$, let μ_x^n denote the n – fold product of μ_x with itself on Ω^n, $X_i(\omega) = \omega_i \in \Omega \subset \mathbb{R}^d$ for $\omega \in \Omega^n$ and let

$$S_n = (S_n^1,\dots,S_n^d) := (X_1 + X_2 + \cdots + X_n)/n,$$

so $S_n : \Omega^n \to \mathbb{R}^d$. The reader is asked to verify (Exercise 20.3) that

$$\int_{\Omega^n} S_n d\mu_x^n := \left(\int_{\Omega^n} S_n^1 d\mu_x^n, \dots, \int_{\Omega^n} S_n^d d\mu_x^n \right) = (x_1, \dots, x_d) = x \tag{20.46}$$

and

$$\int_{\Omega^n} |S_n - x|^2 \, d\mu_x^n = \frac{1}{n} \sum_{i=1}^{d} x_i(1 - x_i) \le \frac{d}{n}. \tag{20.47}$$

From these equations it follows that S_n is concentrating near x as $n \to \infty$, a manifestation of the law of large numbers. Therefore it is reasonable to expect

$$p_n(x) := \int_{\Omega^n} f(S_n) d\mu_x^n \tag{20.48}$$

should approach $f(x)$ as $n \to \infty$. Let $\varepsilon > 0$ be given, $M = \sup\{|f(x)| : x \in [0,1]\}$ and

$$\delta_\varepsilon = \sup\{|f(y) - f(x)| : x, y \in [\mathbf{0}, \mathbf{1}] \text{ and } |y - x| \le \varepsilon\}.$$

By uniform continuity of f on $[\mathbf{0}, \mathbf{1}]$, $\lim_{\varepsilon \downarrow 0} \delta_\varepsilon = 0$. Using these definitions and the fact that $\mu_x^n(\Omega^n) = 1$,

$$\begin{aligned} |f(x) - p_n(x)| &= \left| \int_{\Omega^n} (f(x) - f(S_n)) \, d\mu_x^n \right| \le \int_{\Omega^n} |f(x) - f(S_n)| \, d\mu_x^n \\ &\le \int_{\{|S_n - x| > \varepsilon\}} |f(x) - f(S_n)| \, d\mu_x^n + \int_{\{|S_n - x| \le \varepsilon\}} |f(x) - f(S_n)| \, d\mu_x^n \\ &\le 2M \mu_x^n (|S_n - x| > \varepsilon) + \delta_\varepsilon. \end{aligned} \tag{20.49}$$

By Chebyshev's inequality,

$$\mu_x^n (|S_n - x| > \varepsilon) \le \frac{1}{\varepsilon^2} \int_{\Omega^n} (S_n - x)^2 d\mu_x^n = \frac{d}{n\varepsilon^2},$$

and therefore, Eq. (20.49) yields the estimate

$$\|f - p_n\|_\infty \le \frac{2dM}{n\varepsilon^2} + \delta_\varepsilon$$

and hence

$$\limsup_{n \to \infty} \|f - p_n\|_\infty \le \delta_\varepsilon \to 0 \text{ as } \varepsilon \downarrow 0.$$

This completes the proof since, using Eq. (20.45),

$$p_n(x) = \sum_{\omega \in \Omega^n} f(S_n(\omega)) \mu_x^n(\{\omega\}) = \sum_{\omega \in \Omega^n} f(S_n(\omega)) \prod_{i=1}^{n} \mu_x(\{\omega_i\}),$$

is an nd − degree polynomial in $x \in \mathbb{R}^d$). ■

Exercise 20.3. Verify Eqs. (20.46) and (20.47). This is most easily done using Eqs. (20.43) and (20.44) and Fubini's theorem repeatedly. (Of course Fubini's theorem here is over kill since these are only finite sums after all. Nevertheless it is convenient to use this formulation.)

The second proof requires the next two lemmas.

Lemma 20.30 (Approximate $\delta-$sequences). *Suppose that $\{Q_n\}_{n=1}^{\infty}$ is a sequence of positive functions on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} Q_n(x)\ dx = 1 \text{ and} \tag{20.50}$$

$$\lim_{n\to\infty} \int\limits_{|x|\geq\varepsilon} Q_n(x)dx = 0 \text{ for all } \varepsilon > 0. \tag{20.51}$$

For $f \in BC(\mathbb{R}^d)$, $Q_n * f$ converges to f uniformly on compact subsets of $\mathbb{R}^d$.

Proof. The proof is exactly the same as the proof of Lemma 8.28, it is only necessary to replace $\mathbb{R}$ by $\mathbb{R}^d$ everywhere in the proof. ■

Define

$$Q_n : \mathbb{R}^n \to [0,\infty) \text{ by } Q_n(x) = q_n(x_1)\dots q_n(x_d). \tag{20.52}$$

where q_n is defined in Eq. (8.23).

Lemma 20.31. *The sequence $\{Q_n\}_{n=1}^{\infty}$ is an approximate δ – sequence, i.e. they satisfy Eqs. (20.50) and (20.51).*

Proof. The fact that Q_n integrates to one is an easy consequence of Tonelli's theorem and the fact that q_n integrates to one. Since all norms on $\mathbb{R}^d$ are equivalent, we may assume that $|x| = \max\{|x_i| : i = 1, 2, \dots, d\}$ when proving Eq. (20.51). With this norm

$$\left\{x \in \mathbb{R}^d : |x| \geq \varepsilon\right\} = \cup_{i=1}^d \left\{x \in \mathbb{R}^d : |x_i| \geq \varepsilon\right\}$$

and therefore by Tonelli's theorem,

$$\int\limits_{\{|x|\geq\varepsilon\}} Q_n(x)dx \leq \sum_{i=1}^d \int\limits_{\{|x_i|\geq\varepsilon\}} Q_n(x)dx = d \int\limits_{\{x\in\mathbb{R}|x|\geq\varepsilon\}} q_n(t)dt$$

which tends to zero as $n \to \infty$ by Lemma 8.29. ■

Proof. Proof of Theorem 8.34. Again we assume $f \in C\left(\mathbb{R}^d, \mathbb{C}\right)$ and $f \equiv 0$ on Q_d^c where $Q_d := (0,1)^d$. Let $Q_n(x)$ be defined as in Eq. (20.52). Then by Lemma 20.31 and 20.30, $p_n(x) := (Q_n * F)(x) \to F(x)$ uniformly for $x \in [\mathbf{0}, \mathbf{1}]$ as $n \to \infty$. So to finish the proof it only remains to show $p_n(x)$ is a polynomial when $x \in [\mathbf{0}, \mathbf{1}]$. For $x \in [\mathbf{0}, \mathbf{1}]$,

$$\begin{aligned} p_n(x) &= \int_{\mathbb{R}^d} Q_n(x-y)f(y)dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d \left[c_n^{-1}(1-(x_i-y_i)^2)^n 1_{|x_i-y_i|\leq 1}\right]dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d \left[c_n^{-1}(1-(x_i-y_i)^2)^n\right]dy. \end{aligned}$$

Since the product in the above integrand is a polynomial if $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$, it follows easily that $p_n(x)$ is polynomial in x. ■

20.5 More Spherical Coordinates

In this section we will define spherical coordinates in all dimensions. Along the way we will develop an explicit method for computing surface integrals on spheres. As usual when $n = 2$ define spherical coordinates $(r, \theta) \in (0, \infty) \times [0, 2\pi)$ so that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = T_2(\theta, r).$$

For $n = 3$ we let $x_3 = r\cos\phi_1$ and then

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T_2(\theta, r\sin\phi_1),$$

as can be seen from Figure 20.6, so that

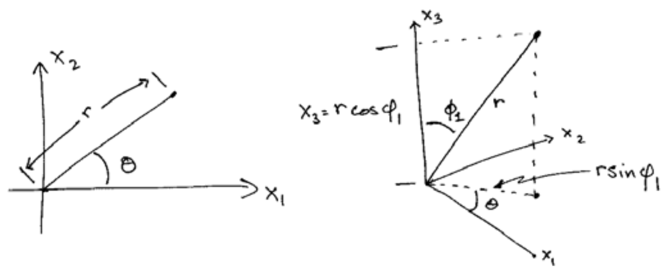


Fig. 20.6. Setting up polar coordinates in two and three dimensions.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} T_2(\theta, r\sin\phi_1) \\ r\cos\phi_1 \end{pmatrix} = \begin{pmatrix} r\sin\phi_1\cos\theta \\ r\sin\phi_1\sin\theta \\ r\cos\phi_1 \end{pmatrix} =: T_3(\theta, \phi_1, r,).$$

We continue to work inductively this way to define

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} T_n(\theta, \phi_1, \dots, \phi_{n-2}, r\sin\phi_{n-1},) \\ r\cos\phi_{n-1} \end{pmatrix} = T_{n+1}(\theta, \phi_1, \dots, \phi_{n-2}, \phi_{n-1}, r).$$

So for example,

$$\begin{aligned}
x_1 &= r\sin\phi_2\sin\phi_1\cos\theta \\
x_2 &= r\sin\phi_2\sin\phi_1\sin\theta \\
x_3 &= r\sin\phi_2\cos\phi_1 \\
x_4 &= r\cos\phi_2
\end{aligned}$$

and more generally,

$$\begin{aligned}
x_1 &= r\sin\phi_{n-2}\dots\sin\phi_2\sin\phi_1\cos\theta \\
x_2 &= r\sin\phi_{n-2}\dots\sin\phi_2\sin\phi_1\sin\theta \\
x_3 &= r\sin\phi_{n-2}\dots\sin\phi_2\cos\phi_1 \\
&\vdots \\
x_{n-2} &= r\sin\phi_{n-2}\sin\phi_{n-3}\cos\phi_{n-4} \\
x_{n-1} &= r\sin\phi_{n-2}\cos\phi_{n-3} \\
x_n &= r\cos\phi_{n-2}.
\end{aligned} \tag{20.53}$$

By the change of variables formula,

$$\begin{aligned}
&\int_{\mathbb{R}^n} f(x)dm(x) \\
&= \int_0^\infty dr \int_{0\le\phi_i\le\pi, 0\le\theta\le 2\pi} d\phi_1\dots d\phi_{n-2}d\theta\,\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r) f(T_n(\theta,\phi_1,\dots,\phi_{n-2},r))
\end{aligned} \tag{20.54}$$

where

$$\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r) := \left|\det T_n'(\theta,\phi_1,\dots,\phi_{n-2},r)\right|.$$

Proposition 20.32. *The Jacobian, Δ_n is given by*

$$\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r) = r^{n-1}\sin^{n-2}\phi_{n-2}\dots\sin^2\phi_2\sin\phi_1. \tag{20.55}$$

If f is a function on rS^{n-1} – the sphere of radius r centered at 0 inside of $\mathbb{R}^n$, then

$$\begin{aligned}
&\int_{rS^{n-1}} f(x)d\sigma(x) = r^{n-1}\int_{S^{n-1}} f(r\omega)d\sigma(\omega) \\
&= \int_{0\le\phi_i\le\pi, 0\le\theta\le 2\pi} f(T_n(\theta,\phi_1,\dots,\phi_{n-2},r))\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r)d\phi_1\dots d\phi_{n-2}d\theta
\end{aligned} \tag{20.56}$$

Proof. We are going to compute Δ_n inductively. Letting $\rho := r\sin\phi_{n-1}$ and writing $\frac{\partial T_n}{\partial\xi}$ for $\frac{\partial T_n}{\partial\xi}(\theta,\phi_1,\dots,\phi_{n-2},\rho)$ we have

$$\begin{aligned}
\Delta_{n+1}(\theta,&\phi_1,\dots,\phi_{n-2},\phi_{n-1},r)\\
&=\left|\begin{bmatrix} \frac{\partial T_n}{\partial\theta} & \frac{\partial T_n}{\partial\phi_1} & \dots & \frac{\partial T_n}{\partial\phi_{n-2}} & \frac{\partial T_n}{\partial\rho} r\cos\phi_{n-1} & \frac{\partial T_n}{\partial\rho}\sin\phi_{n-1} \\ 0 & 0 & \dots & 0 & -r\sin\phi_{n-1} & \cos\phi_{n-1}\end{bmatrix}\right|\\
&= r\left(\cos^2\phi_{n-1}+\sin^2\phi_{n-1}\right)\Delta_n(,\theta,\phi_1,\dots,\phi_{n-2},\rho)\\
&= r\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r\sin\phi_{n-1}),
\end{aligned}$$

i.e.

$$\Delta_{n+1}(\theta,\phi_1,\dots,\phi_{n-2},\phi_{n-1},r) = r\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r\sin\phi_{n-1}). \quad (20.57)$$

To arrive at this result we have expanded the determinant along the bottom row. Staring with $\Delta_2(\theta,r)=r$ already derived in Example 20.23, Eq. (20.57) implies,

$$\begin{aligned}
\Delta_3(\theta,\phi_1,r) &= r\Delta_2(\theta,r\sin\phi_1) = r^2\sin\phi_1\\
\Delta_4(\theta,\phi_1,\phi_2,r) &= r\Delta_3(\theta,\phi_1,r\sin\phi_2) = r^3\sin^2\phi_2\sin\phi_1\\
&\vdots\\
\Delta_n(\theta,\phi_1,\dots,\phi_{n-2},r) &= r^{n-1}\sin^{n-2}\phi_{n-2}\dots\sin^2\phi_2\sin\phi_1
\end{aligned}$$

which proves Eq. (20.55). Eq. (20.56) now follows from Eqs. (50.3), (20.54) and (20.55). ■

As a simple application, Eq. (20.56) implies

$$\begin{aligned}
\sigma(S^{n-1}) &= \int_{0\le\phi_i\le\pi,0\le\theta\le2\pi}\sin^{n-2}\phi_{n-2}\dots\sin^2\phi_2\sin\phi_1 d\phi_1\dots d\phi_{n-2}d\theta\\
&= 2\pi\prod_{k=1}^{n-2}\gamma_k = \sigma(S^{n-2})\gamma_{n-2} \quad (20.58)
\end{aligned}$$

where $\gamma_k := \int_0^\pi \sin^k\phi d\phi$. If $k\ge1$, we have by integration by parts that,

$$\begin{aligned}
\gamma_k &= \int_0^\pi\sin^k\phi d\phi = -\int_0^\pi\sin^{k-1}\phi\ d\cos\phi = 2\delta_{k,1}+(k-1)\int_0^\pi\sin^{k-2}\phi\ \cos^2\phi d\phi\\
&= 2\delta_{k,1}+(k-1)\int_0^\pi\sin^{k-2}\phi\ \left(1-\sin^2\phi\right)d\phi = 2\delta_{k,1}+(k-1)\left[\gamma_{k-2}-\gamma_k\right]
\end{aligned}$$

and hence γ_k satisfies $\gamma_0=\pi$, $\gamma_1=2$ and the recursion relation

$$\gamma_k = \frac{k-1}{k}\gamma_{k-2} \text{ for } k\ge2.$$

Hence we may conclude

$$\gamma_0=\pi,\ \gamma_1=2,\ \gamma_2=\frac12\pi,\ \gamma_3=\frac232,\ \gamma_4=\frac34\frac12\pi,\ \gamma_5=\frac45\frac232,\ \gamma_6=\frac56\frac34\frac12\pi$$

and more generally by induction that

$$\gamma_{2k} = \pi \frac{(2k-1)!!}{(2k)!!} \text{ and } \gamma_{2k+1} = 2\frac{(2k)!!}{(2k+1)!!}.$$

Indeed,

$$\gamma_{2(k+1)+1} = \frac{2k+2}{2k+3}\gamma_{2k+1} = \frac{2k+2}{2k+3}2\frac{(2k)!!}{(2k+1)!!} = 2\frac{[2(k+1)]!!}{(2(k+1)+1)!!}$$

and

$$\gamma_{2(k+1)} = \frac{2k+1}{2k+1}\gamma_{2k} = \frac{2k+1}{2k+2}\pi\frac{(2k-1)!!}{(2k)!!} = \pi\frac{(2k+1)!!}{(2k+2)!!}.$$

The recursion relation in Eq. (20.58) may be written as

$$\sigma(S^n) = \sigma\left(S^{n-1}\right)\gamma_{n-1} \tag{20.59}$$

which combined with $\sigma\left(S^1\right) = 2\pi$ implies

$$\begin{aligned}
\sigma\left(S^1\right) &= 2\pi,\\
\sigma(S^2) &= 2\pi\cdot\gamma_1 = 2\pi\cdot 2,\\
\sigma(S^3) &= 2\pi\cdot 2\cdot\gamma_2 = 2\pi\cdot 2\cdot\frac{1}{2}\pi = \frac{2^2\pi^2}{2!!},\\
\sigma(S^4) &= \frac{2^2\pi^2}{2!!}\cdot\gamma_3 = \frac{2^2\pi^2}{2!!}\cdot 2\frac{2}{3} = \frac{2^3\pi^2}{3!!}\\
\sigma(S^5) &= 2\pi\cdot 2\cdot\frac{1}{2}\pi\cdot\frac{2}{3}2\cdot\frac{3}{4}\frac{1}{2}\pi = \frac{2^3\pi^3}{4!!},\\
\sigma(S^6) &= 2\pi\cdot 2\cdot\frac{1}{2}\pi\cdot\frac{2}{3}2\cdot\frac{3}{4}\frac{1}{2}\pi\cdot\frac{4}{5}\frac{2}{3}2 = \frac{2^4\pi^3}{5!!}
\end{aligned}$$

and more generally that

$$\sigma(S^{2n}) = \frac{2\,(2\pi)^n}{(2n-1)!!} \text{ and } \sigma(S^{2n+1}) = \frac{(2\pi)^{n+1}}{(2n)!!} \tag{20.60}$$

which is verified inductively using Eq. (20.59). Indeed,

$$\sigma(S^{2n+1}) = \sigma(S^{2n})\gamma_{2n} = \frac{2\,(2\pi)^n}{(2n-1)!!}\pi\frac{(2n-1)!!}{(2n)!!} = \frac{(2\pi)^{n+1}}{(2n)!!}$$

and

$$\sigma(S^{(n+1)}) = \sigma(S^{2n+2}) = \sigma(S^{2n+1})\gamma_{2n+1} = \frac{(2\pi)^{n+1}}{(2n)!!}2\frac{(2n)!!}{(2n+1)!!} = \frac{2\,(2\pi)^{n+1}}{(2n+1)!!}.$$

Using

$$(2n)!! = 2n\,(2(n-1))\ldots(2\cdot 1) = 2^n n!$$

we may write $\sigma(S^{2n+1}) = \frac{2\pi^{n+1}}{n!}$ which shows that Eqs. (50.9) and (20.60 in agreement. We may also write the formula in Eq. (20.60) as

$$\sigma(S^n) = \begin{cases} \frac{2(2\pi)^{n/2}}{(n-1)!!} & \text{for } n \text{ even} \\ \frac{(2\pi)^{\frac{n+1}{2}}}{(n-1)!!} & \text{for } n \text{ odd.} \end{cases}$$

20.6 Sard's Theorem

See p. 538 of Taylor and references. Also see Milnor's topology book. Add in the Brower's Fixed point theorem here as well. Also Spivak's calculus on manifolds.

Theorem 20.33. *Let $U \subset_o \mathbb{R}^m$, $f \in C^\infty(U, \mathbb{R}^d)$ and $C := \{x \in U : \operatorname{rank}(f'(x)) < n\}$ be the set of critical points of f. Then the critical values, $f(C)$, is a Borel measurable subset of $\mathbb{R}^d$ of Lebesgue measure* 0.

Remark 20.34. This result clearly extends to manifolds.

For simplicity in the proof given below it will be convenient to use the norm, $|x| := \max_i |x_i|$. Recall that if $f \in C^1(U, \mathbb{R}^d)$ and $p \in U$, then

$$f(p+x) = f(p) + \int_0^1 f'(p+tx)xdt = f(p) + f'(p)x + \int_0^1 [f'(p+tx) - f'(p)]\, xdt$$

so that if

$$R(p,x) := f(p+x) - f(p) - f'(p)x = \int_0^1 [f'(p+tx) - f'(p)]\, xdt$$

we have

$$|R(p,x)| \le |x| \int_0^1 |f'(p+tx) - f'(p)|\, dt = |x|\, \varepsilon(p,x).$$

By uniform continuity, it follows for any compact subset $K \subset U$ that

$$\sup\{|\varepsilon(p,x)| : p \in K \text{ and } |x| \le \delta\} \to 0 \text{ as } \delta \downarrow 0.$$

Proof. Notice that if $x \in U \setminus C$, then $f'(x) : \mathbb{R}^m \to \mathbb{R}^n$ is surjective, which is an open condition, so that $U \setminus C$ is an open subset of U. This shows C is relatively closed in U, i.e. there exists $\tilde{C} \sqsubset \mathbb{R}^m$ such that $C = \tilde{C} \cap U$. Let $K_n \subset U$ be compact subsets of U such that $K_n \uparrow U$, then $K_n \cap C \uparrow C$ and $K_n \cap C = K_n \cap \tilde{C}$ is compact for each n. Therefore, $f(K_n \cap C) \uparrow f(C)$ i.e. $f(C) = \cup_n f(K_n \cap C)$ is a countable union of compact sets and therefore is Borel measurable. Moreover, since $m(f(C)) = \lim_{n\to\infty} m(f(K_n \cap C))$, it suffices to show $m(f(K)) = 0$ for all compact subsets $K \subset C$. Case 1. $(n \le m)$

Let $K = [a, a+\gamma]$ be a cube contained in U and by scaling the domain we may assume $\gamma = (1,1,1,\dots,1)$. For $N \in \mathbb{N}$ and $j \in S_N := \{0,1,\dots,N-1\}^n$ let $K_j = j/N + [a, a+\gamma/N]$ so that $K = \cup_{j\in S_N} K_j$ with $K_j^o \cap K_{j'}^o = \emptyset$ if $j \neq j'$. Let $\{Q_j : j = 1\dots,M\}$ be the collection of those $\{K_j : j \in S_N\}$ which intersect C. For each j, let $p_j \in Q_j \cap C$ and for $x \in Q_j - p_j$ we have

$$f(p_j + x) = f(p_j) + f'(p_j)x + R_j(x)$$

where $|R_j(x)| \le \varepsilon_j(N)/N$ and $\varepsilon(N) := \max_j \varepsilon_j(N) \to 0$ as $N \to \infty$. Now

$$\begin{aligned} m\left(f(Q_j)\right) &= m\left(f(p_j) + \left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \\ &= m\left(\left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \\ &= m\left(O_j\left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \end{aligned} \tag{20.61}$$

where $O_j \in SO(n)$ is chosen so that $O_j f'(p_j)\mathbb{R}^n \subset \mathbb{R}^{m-1} \times \{0\}$. Now $O_j f'(p_j)(Q_j - p_j)$ is contained in $\Gamma \times \{0\}$ where $\Gamma \subset \mathbb{R}^{m-1}$ is a cube centered at $0 \in \mathbb{R}^{m-1}$ with side length at most $2\,|f'(p_j)|/N \le 2M/N$ where $M = \max_{p\in K} |f'(p)|$. It now follows that $O_j\left(f'(p_j) + R_j\right)(Q_j - p_j)$ is contained the set of all points within $\varepsilon(N)/N$ of $\Gamma \times \{0\}$ and in particular

$$O_j\left(f'(p_j) + R_j\right)(Q_j - p_j) \subset \left(1 + \varepsilon(N)/N\right)\Gamma \times [\varepsilon(N)/N, \varepsilon(N)/N].$$

From this inclusion and Eq. (20.61) it follows that

$$\begin{aligned} m\left(f(Q_j)\right) &\le \left[2\frac{M}{N}\left(1 + \varepsilon(N)/N\right)\right]^{m-1} 2\varepsilon(N)/N \\ &= 2^m M^{m-1}\left[\left(1 + \varepsilon(N)/N\right)\right]^{m-1}\varepsilon(N)\frac{1}{N^m} \end{aligned}$$

and therefore,

$$\begin{aligned} m\left(f(C\cap K)\right) &\le \sum_j m\left(f(Q_j)\right) \le N^n 2^m M^{m-1}\left[\left(1 + \varepsilon(N)/N\right)\right]^{m-1}\varepsilon(N)\frac{1}{N^m} \\ &= 2^n M^{n-1}\left[\left(1 + \varepsilon(N)/N\right)\right]^{n-1}\varepsilon(N)\frac{1}{N^{m-n}} \to 0 \text{ as } N \to \infty \end{aligned}$$

since $m \ge n$. This proves the easy case since we may write U as a countable union of cubes K as above. **Remark**. The case $(m < n)$ also follows from the case $m = n$ as follows. When $m < n$, $C = U$ and we must show $m(f(U)) = 0$. Letting $F : U \times \mathbb{R}^{n-m} \to \mathbb{R}^n$ be the map $F(x,y) = f(x)$. Then $F'(x,y)(v,w) = f'(x)v$, and hence $C_F := U \times \mathbb{R}^{n-m}$. So if the assertion holds for $m = n$ we have

$$m(f(U)) = m(F(U \times \mathbb{R}^{n-m})) = 0.$$

Case 2. $(m > n)$ This is the hard case and the case we will need in the co-area formula to be proved later. Here I will follow the proof in Milnor. Let

$$C_i := \{x \in U : \partial^\alpha f(x) = 0 \text{ when } |\alpha| \le i\}$$

so that $C \supset C_1 \supset C_2 \supset C_3 \supset \dots$. The proof is by induction on n and goes by the following steps:

1. $m(f(C \setminus C_1)) = 0$.
2. $m(f(C_i \setminus C_{i+1})) = 0$ for all $i \geq 1$.
3. $m(f(C_i)) = 0$ for all i sufficiently large.

Step 1. If $m = 1$, there is nothing to prove since $C = C_1$ so we may assume $m \geq 2$. Suppose that $x \in C \setminus C_1$, then $f'(p) \neq 0$ and so by reordering the components of x and $f(p)$ if necessary we may assume that $\partial_1 f_1(p) \neq 0$ where we are writing $\partial f(p)/\partial x_i$ as $\partial_i f(p)$. The map $h(x) := (f_1(x), x_2, \ldots, x_n)$ has differential

$$h'(p) = \begin{bmatrix} \partial_1 f_1(p) & \partial_2 f_1(p) & \ldots & \partial_n f_1(p) \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which is not singular. So by the implicit function theorem, there exists there exists $V \in \tau_p$ such that $h : V \to h(V) \in \tau_{h(p)}$ is a diffeomorphism and in particular $\partial f_1(x)/\partial x_1 \neq 0$ for $x \in V$ and hence $V \subset U \setminus C_1$. Consider the map $g := f \circ h^{-1} : V' := h(V) \to \mathbb{R}^m$, which satisfies

$$(f_1(x), f_2(x), \ldots, f_m(x)) = f(x) = g(h(x)) = g((f_1(x), x_2, \ldots, x_n))$$

which implies $g(t, y) = (t, u(t, y))$ for $(t, y) \in V' := h(V) \in \tau_{h(p)}$, see Figure 20.7 below where $p = \bar{x}$ and $m = p$. Since

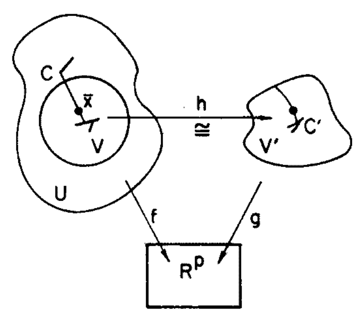


Fig. 20.7. Making a change of variable so as to apply induction.

$$g'(t, y) = \begin{bmatrix} 1 & 0 \\ \partial_t u(t, y) & \partial_y u(t, y) \end{bmatrix}$$

it follows that (t, y) is a critical point of g iff $y \in C'_t$ – the set of critical points of $y \to u(t, y)$. Since h is a diffeomorphism we have $C' := h(C \cap V)$ are the critical points of g in V' and

$$f(C\cap V)=g(C')=\cup_t\left[\{t\}\times u_t(C'_t)\right].$$

By the induction hypothesis, $m_{m-1}(u_t(C'_t))=0$ for all t, and therefore by Fubini's theorem,

$$m(f(C\cap V))=\int_{\mathbb{R}}m_{m-1}(u_t(C'_t))1_{V'_t\neq\emptyset}dt=0.$$

Since $C\setminus C_1$ may be covered by a countable collection of open sets V as above, it follows that $m(f(C\setminus C_1))=0$. **Step 2.** Suppose that $p\in C_k\setminus C_{k+1}$, then there is an α such that $|\alpha|=k+1$ such that $\partial^\alpha f(p)=0$ while $\partial^\beta f(p)=0$ for all $|\beta|\le k$. Again by permuting coordinates we may assume that $\alpha_1\neq 0$ and $\partial^\alpha f_1(p)\neq 0$. Let $w(x):=\partial^{\alpha-e_1}f_1(x)$, then $w(p)=0$ while $\partial_1 w(p)\neq 0$. So again the implicit function theorem there exists $V\in\tau_p$ such that $h(x):=(w(x),x_2,\ldots,x_n)$ maps $V\to V':=h(V)\in\tau_{h(p)}$ in a diffeomorphic way and in particular $\partial_1 w(x)\neq 0$ on V so that $V\subset U\setminus C_{k+1}$. As before, let $g:=f\circ h^{-1}$ and notice that $C'_k:=h(C_k\cap V)\subset\{0\}\times\mathbb{R}^{n-1}$ and

$$f(C_k\cap V)=g(C'_k)=\bar g\left(C'_k\right)$$

where $\bar g:=g|_{(\{0\}\times\mathbb{R}^{n-1})\cap V'}$. Clearly C'_k is contained in the critical points of $\bar g$, and therefore, by induction

$$0=m(\bar g(C'_k))=m(f(C_k\cap V)).$$

Since $C_k\setminus C_{k+1}$ is covered by a countable collection of such open sets, it follows that

$$m(f(C_k\setminus C_{k+1}))=0\text{ for all }k\ge 1.$$

Step 3. Suppose that Q is a closed cube with edge length δ contained in U and $k>n/m-1$. We will show $m(f(Q\cap C_k))=0$ and since Q is arbitrary it will follows that $m(f(C_k))=0$ as desired. By Taylor's theorem with (integral) remainder, it follows for $x\in Q\cap C_k$ and h such that $x+h\in Q$ that

$$f(x+h)=f(x)+R(x,h)$$

where

$$|R(x,h)|\le c\left\|h\right\|^{k+1}$$

where $c=c(Q,k)$. Now subdivide Q into r^n cubes of edge size δ/r and let Q' be one of the cubes in this subdivision such that $Q'\cap C_k\neq\emptyset$ and let $x\in Q'\cap C_k$. It then follows that $f(Q')$ is contained in a cube centered at $f(x)\in\mathbb{R}^m$ with side length at most $2c\left(\delta/r\right)^{k+1}$ and hence volume at most $\left(2c\right)^m\left(\delta/r\right)^{m(k+1)}$. Therefore, $f(Q\cap C_k)$ is contained in the union of at most r^n cubes of volume $\left(2c\right)^m\left(\delta/r\right)^{m(k+1)}$ and hence $m$each

$$m\left(f(Q\cap C_k)\right)\le\left(2c\right)^m\left(\delta/r\right)^{m(k+1)}r^n=\left(2c\right)^m\delta^{m(k+1)}r^{n-m(k+1)}\to 0\text{ as }r\uparrow\infty$$

provided that $n-m(k+1)<0$, i.e. provided $k>n/m-1$. ∎

20.7 Exercises

Exercise 20.4. Prove Theorem 20.12. Suggestion, to get started define

$$\pi(A) := \int_{X_1} d\mu(x_1) \dots \int_{X_n} d\mu(x_n)\, 1_A(x_1, \dots, x_n)$$

and then show Eq. (20.18) holds. Use the case of two factors as the model of your proof.

Exercise 20.5. Let $(X_j, \mathcal{M}_j, \mu_j)$ for $j = 1, 2, 3$ be σ – finite measure spaces. Let $F : (X_1 \times X_2) \times X_3 \to X_1 \times X_2 \times X_3$ be defined by

$$F((x_1, x_2), x_3) = (x_1, x_2, x_3).$$

1. Show F is $((\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$ – measurable and F^{-1} is $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3)$ – measurable. That is
$$F : ((X_1 \times X_2) \times X_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3) \to (X_1 \times X_2 \times X_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$$
is a "measure theoretic isomorphism."
2. Let $\pi := F_* [(\mu_1 \otimes \mu_2) \otimes \mu_3]$, i.e. $\pi(A) = [(\mu_1 \otimes \mu_2) \otimes \mu_3](F^{-1}(A))$ for all $A \in \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$. Then π is the unique measure on $\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$ such that
$$\pi(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3)$$
for all $A_i \in \mathcal{M}_i$. We will write $\pi := \mu_1 \otimes \mu_2 \otimes \mu_3$.
3. Let $f : X_1 \times X_2 \times X_3 \to [0, \infty]$ be a $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function. Verify the identity,
$$\int_{X_1 \times X_2 \times X_3} f d\pi = \int_{X_3} d\mu_3(x_3) \int_{X_2} d\mu_2(x_2) \int_{X_1} d\mu_1(x_1) f(x_1, x_2, x_3),$$
makes sense and is correct.
4. (Optional.) Also show the above identity holds for any one of the six possible orderings of the iterated integrals.

Exercise 20.6. Prove the second assertion of Theorem 20.18. That is show m^d is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0, 1]^d) = 1$. **Hint:** Look at the proof of Theorem 19.10.

Exercise 20.7. (Part of Folland Problem 2.46 on p. 69.) Let $X = [0, 1]$, $\mathcal{M} = \mathcal{B}_{[0,1]}$ be the Borel σ – field on X, m be Lebesgue measure on $[0, 1]$ and ν be counting measure, $\nu(A) = \#(A)$. Finally let $D = \{(x, x) \in X^2 : x \in X\}$ be the diagonal in X^2. Show

$$\int_X \left[\int_X 1_D(x, y) d\nu(y) \right] dm(x) \neq \int_X \left[\int_X 1_D(x, y) dm(x) \right] d\nu(y)$$

by explicitly computing both sides of this equation.

Exercise 20.8. Folland Problem 2.48 on p. 69. (Counter example related to Fubini Theorem involving counting measures.)

Exercise 20.9. Folland Problem 2.50 on p. 69 pertaining to area under a curve. (Note the $\mathcal{M} \times \mathcal{B}_{\mathbb{R}}$ should be $\mathcal{M} \otimes \mathcal{B}_{\bar{\mathbb{R}}}$ in this problem.)

Exercise 20.10. Folland Problem 2.55 on p. 77. (Explicit integrations.)

Exercise 20.11. Folland Problem 2.56 on p. 77. Let $f \in L^1((0,a), dm)$, $g(x) = \int_x^a \frac{f(t)}{t} dt$ for $x \in (0,a)$, show $g \in L^1((0,a), dm)$ and

$$\int_0^a g(x)dx = \int_0^a f(t)dt.$$

Exercise 20.12. Show $\int_0^\infty \left|\frac{\sin x}{x}\right| dm(x) = \infty$. So $\frac{\sin x}{x} \notin L^1([0,\infty), m)$ and $\int_0^\infty \frac{\sin x}{x} dm(x)$ is not defined as a Lebesgue integral.

Exercise 20.13. Folland Problem 2.57 on p. 77.

Exercise 20.14. Folland Problem 2.58 on p. 77.

Exercise 20.15. Folland Problem 2.60 on p. 77. Properties of the Γ – function.

Exercise 20.16. Folland Problem 2.61 on p. 77. Fractional integration.

Exercise 20.17. Folland Problem 2.62 on p. 80. Rotation invariance of surface measure on S^{n-1}.

Exercise 20.18. Folland Problem 2.64 on p. 80. On the integrability of $|x|^a \left|\log |x|\right|^b$ for x near 0 and x near ∞ in $\mathbb{R}^n$.

21

L^p-spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space and for $0 < p < \infty$ and a measurable function $f : X \to \mathbb{C}$ let

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}. \tag{21.1}$$

When $p = \infty$, let

$$\|f\|_\infty = \inf \{ a \geq 0 : \mu(|f| > a) = 0 \} \tag{21.2}$$

For $0 < p \leq \infty$, let

$$L^p(X, \mathcal{M}, \mu) = \{ f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_p < \infty \} / \sim$$

where $f \sim g$ iff $f = g$ a.e. Notice that $\|f - g\|_p = 0$ iff $f \sim g$ and if $f \sim g$ then $\|f\|_p = \|g\|_p$. In general we will (by abuse of notation) use f to denote both the function f and the equivalence class containing f.

Remark 21.1. Suppose that $\|f\|_\infty \leq M$, then for all $a > M$, $\mu(|f| > a) = 0$ and therefore $\mu(|f| > M) = \lim_{n\to\infty} \mu(|f| > M + 1/n) = 0$, i.e. $|f(x)| \leq M$ for μ - a.e. x. Conversely, if $|f| \leq M$ a.e. and $a > M$ then $\mu(|f| > a) = 0$ and hence $\|f\|_\infty \leq M$. This leads to the identity:

$$\|f\|_\infty = \inf \{ a \geq 0 : |f(x)| \leq a \text{ for } \mu - \text{a.e. } x \}.$$

The next theorem is a generalization Theorem 5.6 to general integrals and the proof is essentially identical to the proof of Theorem 5.6.

Theorem 21.2 (Hölder's inequality). *Suppose that $1 \leq p \leq \infty$ and $q := \frac{p}{p-1}$, or equivalently $p^{-1} + q^{-1} = 1$. If f and g are measurable functions then*

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q. \tag{21.3}$$

Assuming $p \in (1, \infty)$ and $\|f\|_p \cdot \|g\|_q < \infty$, equality holds in Eq. (21.3) iff $|f|^p$ and $|g|^q$ are linearly dependent as elements of L^1 which happens iff

$$|g|^q \|f\|_p^p = \|g\|_q^q \, |f|^p \quad a.e. \tag{21.4}$$

Proof. The cases where $\|f\|_q = 0$ or ∞ or $\|g\|_p = 0$ or ∞ are easy to deal with and are left to the reader. So we will now assume that $0 < \|f\|_q, \|g\|_p < \infty$. Let $s = |f| / \|f\|_p$ and $t = |g|/\|g\|_q$ then Lemma 5.5 implies

$$\frac{|fg|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \frac{|f|^p}{\|f\|_p} + \frac{1}{q} \frac{|g|^q}{\|g\|^q} \tag{21.5}$$

with equality iff $|g/\|g\|_q| = |f|^{p-1} / \|f\|_p^{(p-1)} = |f|^{p/q} / \|f\|_p^{p/q}$, i.e. $|g|^q \|f\|_p^p = \|g\|_q^q |f|^p$. Integrating Eq. (21.5) implies

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

with equality iff Eq. (21.4) holds. The proof is finished since it is easily checked that equality holds in Eq. (21.3) when $|f|^p = c\,|g|^q$ of $|g|^q = c\,|f|^p$ for some constant c. ■

The following corollary is an easy extension of Hölder's inequality.

Corollary 21.3. *Suppose that $f_i : X \to \mathbb{C}$ are measurable functions for $i = 1, \dots, n$ and $p_1, \dots, p_n$ and r are positive numbers such that $\sum_{i=1}^n p_i^{-1} = r^{-1}$, then*

$$\left\| \prod_{i=1}^n f_i \right\|_r \leq \prod_{i=1}^n \|f_i\|_{p_i} \ \ \textit{where} \ \sum_{i=1}^n p_i^{-1} = r^{-1}.$$

Proof. To prove this inequality, start with $n = 2$, then for any $p \in [1, \infty]$,

$$\|fg\|_r^r = \int_X |f|^r \, |g|^r \, d\mu \leq \|f^r\|_p \, \|g^r\|_{p^*}$$

where $p^* = \frac{p}{p-1}$ is the conjugate exponent. Let $p_1 = pr$ and $p_2 = p^* r$ so that $p_1^{-1} + p_2^{-1} = r^{-1}$ as desired. Then the previous equation states that

$$\|fg\|_r \leq \|f\|_{p_1} \|g\|_{p_2}$$

as desired. The general case is now proved by induction. Indeed,

$$\left\| \prod_{i=1}^{n+1} f_i \right\|_r = \left\| \prod_{i=1}^{n} f_i \cdot f_{n+1} \right\|_r \leq \left\| \prod_{i=1}^{n} f_i \right\|_q \|f_{n+1}\|_{p_{n+1}}$$

where $q^{-1} + p_{n+1}^{-1} = r^{-1}$. Since $\sum_{i=1}^n p_i^{-1} = q^{-1}$, we may now use the induction hypothesis to conclude

$$\left\| \prod_{i=1}^{n} f_i \right\|_q \leq \prod_{i=1}^n \|f_i\|_{p_i},$$

which combined with the previous displayed equation proves the generalized form of Holder's inequality. ■

Theorem 21.4 (Minkowski's Inequality). *If $1 \le p \le \infty$ and $f, g \in L^p$ then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{21.6}$$

Moreover, assuming f and g are not identically zero, equality holds in Eq. (21.6) iff $\operatorname{sgn}(f) \stackrel{\circ}{=} \operatorname{sgn}(g)$ a.e. (see the notation in Definition 5.7) when $p = 1$ and $f = cg$ a.e. for some $c > 0$ for $p \in (1, \infty)$.

Proof. When $p = \infty$, $|f| \le \|f\|_\infty$ a.e. and $|g| \le \|g\|_\infty$ a.e. so that $|f+g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$ a.e. and therefore

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty .$$

When $p < \infty$,

$$|f+g|^p \le (2\max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \le 2^p (|f|^p + |g|^p),$$

$$\|f+g\|_p^p \le 2^p \left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

In case $p = 1$,

$$\|f+g\|_1 = \int_X |f+g| d\mu \le \int_X |f|\, d\mu + \int_X |g| d\mu$$

with equality iff $|f| + |g| = |f+g|$ a.e. which happens iff $\operatorname{sgn}(f) \stackrel{\circ}{=} \operatorname{sgn}(g)$ a.e. In case $p \in (1, \infty)$, we may assume $\|f+g\|_p$, $\|f\|_p$ and $\|g\|_p$ are all positive since otherwise the theorem is easily verified. Now

$$|f+g|^p = |f+g||f+g|^{p-1} \le (|f| + |g|)|f+g|^{p-1}$$

with equality iff $\operatorname{sgn}(f) \stackrel{\circ}{=} \operatorname{sgn}(g)$. Integrating this equation and applying Holder's inequality with $q = p/(p-1)$ gives

$$\begin{aligned} \int_X |f+g|^p d\mu &\le \int_X |f| \ |f+g|^{p-1} d\mu + \int_X |g| \ |f+g|^{p-1} d\mu \\ &\le (\|f\|_p + \|g\|_p) \| \, |f+g|^{p-1} \, \|_q \end{aligned} \tag{21.7}$$

with equality iff

$$\begin{aligned} &\operatorname{sgn}(f) \stackrel{\circ}{=} \operatorname{sgn}(g) \text{ and} \\ &\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f+g|^p}{\|f+g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p \text{ a.e.} \end{aligned} \tag{21.8}$$

Therefore

$$\||f+g|^{p-1}\|_q^q = \int_X (|f+g|^{p-1})^q d\mu = \int_X |f+g|^p d\mu. \tag{21.9}$$

Combining Eqs. (21.7) and (21.9) implies

$$\|f+g\|_p^p \le \|f\|_p \|f+g\|_p^{p/q} + \|g\|_p \|f+g\|_p^{p/q} \tag{21.10}$$

with equality iff Eq. (21.8) holds which happens iff $f = cg$ a.e. with $c > 0$. Solving for $\|f+g\|_p$ in Eq. (21.10) gives Eq. (21.6). ■

The next theorem gives another example of using Hölder's inequality

Theorem 21.5. *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ – finite measure spaces, $p \in [1, \infty]$, $q = p/(p-1)$ and $k : X \times Y \to \mathbb{C}$ be a $\mathcal{M} \otimes \mathcal{N}$ – measurable function. Assume there exist finite constants C_1 and C_2 such that*

$$\int_X |k(x,y)|\, d\mu(x) \le C_1 \text{ for } \nu \text{ a.e. } y \text{ and}$$
$$\int_Y |k(x,y)|\, d\nu(y) \le C_2 \text{ for } \mu \text{ a.e. } x.$$

If $f \in L^p(\nu)$, then

$$\int_Y |k(x,y) f(y)|\, d\nu(y) < \infty \text{ for } \mu \text{ – a.e. } x,$$

$x \to Kf(x) := \int_Y k(x,y) f(y) d\nu(y) \in L^p(\mu)$ and

$$\|Kf\|_{L^p(\mu)} \le C_1^{1/p} C_2^{1/q} \|f\|_{L^p(\nu)} \tag{21.11}$$

Proof. Suppose $p \in (1, \infty)$ to begin with and let $q = p/(p-1)$, then by Hölder's inequality,

$$\begin{aligned}
\int_Y |k(x,y) f(y)|\, d\nu(y) &= \int_Y |k(x,y)|^{1/q} |k(x,y)|^{1/p} |f(y)|\, d\nu(y) \\
&\le \left[\int_Y |k(x,y)|\, d\nu(y)\right]^{1/q} \left[\int_Y |k(x,y)|\, |f(y)|^p\, d\nu(y)\right]^{1/p} \\
&\le C_2^{1/q} \left[\int_Y |k(x,y)|\, |f(y)|^p\, d\nu(y)\right]^{1/p}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left\|\int_Y |k(\cdot,y) f(y)|\, d\nu(y)\right\|_{L^p(\mu)}^p &= \int_X d\mu(x) \left[\int_Y |k(x,y) f(y)|\, d\nu(y)\right]^p \\
&\le C_2^{p/q} \int_X d\mu(x) \int_Y d\nu(y)\, |k(x,y)|\, |f(y)|^p \\
&= C_2^{p/q} \int_Y d\nu(y)\, |f(y)|^p \int_X d\mu(x)\, |k(x,y)| \\
&\le C_2^{p/q} C_1 \int_Y d\nu(y)\, |f(y)|^p = C_2^{p/q} C_1 \|f\|_{L^p(\nu)}^p,
\end{aligned}$$

wherein we used Tonelli's theorem in third line. From this it follows that $\int_Y |k(x,y)f(y)|\, d\nu(y) < \infty$ for μ - a.e. x,

$$x \to Kf(x) := \int_Y k(x,y)f(y)d\nu(y) \in L^p(\mu)$$

and that Eq. (21.11) holds.

Similarly if $p = \infty$,

$$\int_Y |k(x,y)f(y)|\, d\nu(y) \le \|f\|_{L^\infty(\nu)} \cdot \int_Y |k(x,y)|\, d\nu(y) \le C_2 \|f\|_{L^\infty(\nu)} \text{ for } \mu - \text{a.e. } x.$$

so that $\|Kf\|_{L^\infty(\mu)} \le C_2 \|f\|_{L^\infty(\nu)}$. If $p = 1$, then

$$\begin{aligned}\int_X d\mu(x) \int_Y d\nu(y)\, |k(x,y)f(y)| &= \int_Y d\nu(y)\, |f(y)| \int_X d\mu(x)\, |k(x,y)| \\ &\le C_1 \int_Y d\nu(y)\, |f(y)|\end{aligned}$$

which shows $\|Kf\|_{L^1(\mu)} \le C_1 \|f\|_{L^1(\nu)}$. ■

21.1 Jensen's Inequality

Definition 21.6. *A function $\phi : (a,b) \to \mathbb{R}$ is convex if for all $a < x_0 < x_1 < b$ and $t \in [0,1]$ $\phi(x_t) \le t\phi(x_1) + (1-t)\phi(x_0)$ where $x_t = tx_1 + (1-t)x_0$.*

Example 21.7. The functions $\exp(x)$ and $-\log(x)$ are convex and x^p is convex iff $p \ge 1$ as follows from Corollary 21.9 below which in part states that any $\phi \in C^2((a,b), \mathbb{R})$ such that $\phi'' \ge 0$ is convex.

The following Proposition is clearly motivated by Figure 21.1.

Proposition 21.8. *Suppose $\phi : (a,b) \to \mathbb{R}$ is a convex function, then*

1. *For all $u, v, w, z \in (a,b)$ such that $u < z$, $w \in [u, z)$ and $v \in (u, z]$,*
$$\frac{\phi(v) - \phi(u)}{v - u} \le \frac{\phi(z) - \phi(w)}{z - w}. \tag{21.12}$$
2. *For each $c \in (a,b)$, the right and left sided derivatives $\phi'_\pm(c)$ exists in $\mathbb{R}$ and if $a < u < v < b$, then $\phi'_+(u) \le \phi'_-(v) \le \phi'_+(v)$.*
3. *The function ϕ is continuous.*
4. *For all $t \in (a,b)$ and $\beta \in [\phi'_-(t), \phi'_+(t)]$, $\phi(x) \ge \phi(t) + \beta(x - t)$ for all $x \in (a,b)$. In particular,*
$$\phi(x) \ge \phi(t) + \phi'_-(t)(x - t) \text{ for all } x, t \in (a,b).$$

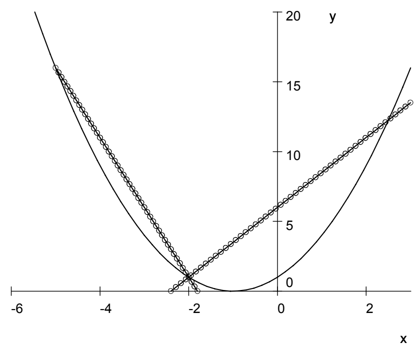


Fig. 21.1. A convex function along with two cords corresponding to $x_0 = -2$ and $x_1 = 4$ and $x_0 = -5$ and $x_1 = -2$.

Proof. 1a) Suppose first that $u < v = w < z$, in which case Eq. (21.12) is equivalent to

$$(\phi(v) - \phi(u))(z - v) \le (\phi(z) - \phi(v))(v - u)$$

which after solving for $\phi(v)$ is equivalent to the following equations holding:

$$\phi(v) \le \phi(z)\frac{v - u}{z - u} + \phi(u)\frac{z - v}{z - u}.$$

But this last equation states that $\phi(v) \le \phi(z)t + \phi(u)(1 - t)$ where $t = \frac{v-u}{z-u}$ and $v = tz + (1 - t)u$ and hence is valid by the definition of ϕ being convex. 1b) Now assume $u = w < v < z$, in which case Eq. (21.12) is equivalent to

$$(\phi(v) - \phi(u))(z - u) \le (\phi(z) - \phi(u))(v - u)$$

which after solving for $\phi(v)$ is equivalent to

$$\phi(v)(z - u) \le \phi(z)(v - u) + \phi(u)(z - v)$$

which is equivalent to

$$\phi(v) \le \phi(z)\frac{v - u}{z - u} + \phi(u)\frac{z - v}{z - u}.$$

Again this equation is valid by the convexity of ϕ. 1c) $u < w < v = z$, in which case Eq. (21.12) is equivalent to

$$(\phi(z) - \phi(u))(z - w) \le (\phi(z) - \phi(w))(z - u)$$

and this is equivalent to the inequality,

$$\phi(w) \leq \phi(z)\frac{w-u}{z-u} + \phi(u)\frac{z-w}{z-u}$$

which again is true by the convexity of ϕ. 1) General case. If $u < w < v < z$, then by 1a-1c)

$$\frac{\phi(z)-\phi(w)}{z-w} \geq \frac{\phi(v)-\phi(w)}{v-w} \geq \frac{\phi(v)-\phi(u)}{v-u}$$

and if $u < v < w < z$

$$\frac{\phi(z)-\phi(w)}{z-w} \geq \frac{\phi(w)-\phi(v)}{w-v} \geq \frac{\phi(w)-\phi(u)}{w-u}.$$

We have now taken care of all possible cases. 2) On the set $a < w < z < b$, Eq. (21.12) shows that $(\phi(z)-\phi(w))\,/\,(z-w)$ is a decreasing function in w and an increasing function in z and therefore $\phi'_{\pm}(x)$ exists for all $x \in (a,b)$. Also from Eq. (21.12) we learn that

$$\phi'_+(u) \leq \frac{\phi(z)-\phi(w)}{z-w} \text{ for all } a < u < w < z < b, \tag{21.13}$$

$$\frac{\phi(v)-\phi(u)}{v-u} \leq \phi'_-(z) \text{ for all } a < u < v < z < b, \tag{21.14}$$

and letting $w \uparrow z$ in the first equation also implies that

$$\phi'_+(u) \leq \phi'_-(z) \text{ for all } a < u < z < b.$$

The inequality, $\phi'_-(z) \leq \phi'_+(z)$, is also an easy consequence of Eq. (21.12). 3) Since $\phi(x)$ has both left and right finite derivatives, it follows that ϕ is continuous. (For an alternative proof, see Rudin.) 4) Given t, let $\beta \in [\phi'_-(t), \phi'_+(t)]$, then by Eqs. (21.13) and (21.14),

$$\frac{\phi(t)-\phi(u)}{t-u} \leq \phi'_-(t) \leq \beta \leq \phi'_+(t) \leq \frac{\phi(z)-\phi(t)}{z-t}$$

for all $a < u < t < z < b$. Item 4. now follows. ∎

Corollary 21.9. *Suppose $\phi : (a,b) \to \mathbb{R}$ is differential then ϕ is convex iff ϕ' is non decreasing. In particular if $\phi \in C^2(a,b)$ then ϕ is convex iff $\phi'' \geq 0$.*

Proof. By Proposition 21.8, if ϕ is convex then ϕ' is non-decreasing. Conversely if ϕ' is increasing then by the mean value theorem,

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} = \phi'(\xi_1) \text{ for some } \xi_1 \in (c, x_1)$$

and

$$\frac{\phi(c)-\phi(x_0)}{c-x_0} = \phi'(\xi_2) \text{ for some } \xi_2 \in (x_0, c).$$

Hence

$$\frac{\phi(x_1) - \phi(c)}{x_1 - c} \geq \frac{\phi(c) - \phi(x_0)}{c - x_0}$$

for all $x_0 < c < x_1$. Solving this inequality for $\phi(c)$ gives

$$\phi(c) \leq \frac{c - x_0}{x_1 - x_0}\phi(x_1) + \frac{x_1 - c}{x_1 - x_0}\phi(x_0)$$

showing ϕ is convex. ■

Theorem 21.10 (Jensen's Inequality). *Suppose that $(X, \mathcal{M}, \mu)$ is a probability space, i.e. μ is a positive measure and $\mu(X) = 1$. Also suppose that $f \in L^1(\mu)$, $f : X \to (a, b)$, and $\phi : (a, b) \to \mathbb{R}$ is a convex function. Then*

$$\phi\left(\int_X f d\mu\right) \leq \int_X \phi(f) d\mu$$

where if $\phi \circ f \notin L^1(\mu)$, then $\phi \circ f$ is integrable in the extended sense and $\int_X \phi(f) d\mu = \infty$.

Proof. Let $t = \int_X f d\mu \in (a, b)$ and let $\beta \in \mathbb{R}$ be such that $\phi(s) - \phi(t) \geq \beta(s-t)$ for all $s \in (a, b)$. Then integrating the inequality, $\phi(f) - \phi(t) \geq \beta(f - t)$, implies that

$$0 \leq \int_X \phi(f) d\mu - \phi(t) = \int_X \phi(f) d\mu - \phi(\int_X f d\mu).$$

Moreover, if $\phi(f)$ is not integrable, then $\phi(f) \geq \phi(t) + \beta(f - t)$ which shows that negative part of $\phi(f)$ is integrable. Therefore, $\int_X \phi(f) d\mu = \infty$ in this case. ■

Example 21.11. The convex functions in Example 21.7 lead to the following inequalities,

$$\exp\left(\int_X f d\mu\right) \leq \int_X e^f d\mu, \tag{21.15}$$

$$\int_X \log(|f|) d\mu \leq \log\left(\int_X |f| \, d\mu\right)$$

and for $p \geq 1$,

$$\left|\int_X f d\mu\right|^p \leq \left(\int_X |f| \, d\mu\right)^p \leq \int_X |f|^p \, d\mu.$$

The last equation may also easily be derived using Hölder's inequality. As a special case of the first equation, we get another proof of Lemma 5.5. Indeed, more generally, suppose $p_i, s_i > 0$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^n \frac{1}{p_i} = 1$, then

$$s_1 \dots s_n = e^{\sum_{i=1}^n \ln s_i} = e^{\sum_{i=1}^n \frac{1}{p_i} \ln s_i^{p_i}} \leq \sum_{i=1}^n \frac{1}{p_i} e^{\ln s_i^{p_i}} = \sum_{i=1}^n \frac{s_i^{p_i}}{p_i} \tag{21.16}$$

where the inequality follows from Eq. (21.15) with $X = \{1, 2, \dots, n\}$, $\mu = \sum_{i=1}^n \frac{1}{p_i} \delta_i$ and $f(i) := \ln s_i^{p_i}$. Of course Eq. (21.16) may be proved directly using the convexity of the exponential function.

21.2 Modes of Convergence

As usual let $(X, \mathcal{M}, \mu)$ be a fixed measure space, assume $1 \leq p \leq \infty$ and let $\{f_n\}_{n=1}^\infty \cup \{f\}$ be a collection of complex valued measurable functions on X. We have the following notions of convergence and Cauchy sequences.

Definition 21.12. *1. $f_n \to f$ a.e. if there is a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $\lim_{n\to\infty} 1_{E^c} f_n = 1_{E^c} f$.*
2. $f_n \to f$ in μ – measure if $\lim_{n\to\infty} \mu(|f_n - f| > \varepsilon) = 0$ for all $\varepsilon > 0$. We will abbreviate this by saying $f_n \to f$ in L^0 or by $f_n \xrightarrow{\mu} f$.
3. $f_n \to f$ in L^p iff $f \in L^p$ and $f_n \in L^p$ for all n, and $\lim_{n\to\infty} \|f_n - f\|_p = 0$.

Definition 21.13. *1. $\{f_n\}$ is a.e. Cauchy if there is a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ and$\{1_{E^c} f_n\}$ is a pointwise Cauchy sequences.*
2. $\{f_n\}$ is Cauchy in μ – measure (or L^0 – Cauchy) if $\lim_{m,n\to\infty} \mu(|f_n - f_m| > \varepsilon) = 0$ for all $\varepsilon > 0$.
3. $\{f_n\}$ is Cauchy in L^p if $\lim_{m,n\to\infty} \|f_n - f_m\|_p = 0$.

Lemma 21.14 (Chebyshev's inequality again). *Let $p \in [1, \infty)$ and $f \in L^p$, then*

$$\mu(|f| \geq \varepsilon) \leq \frac{1}{\varepsilon^p} \|f\|_p^p \text{ for all } \varepsilon > 0.$$

In particular if $\{f_n\} \subset L^p$ is L^p – convergent (Cauchy) then $\{f_n\}$ is also convergent (Cauchy) in measure.

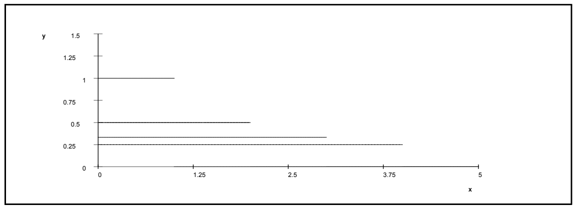
Proof. By Chebyshev's inequality (19.11),

$$\mu(|f| \geq \varepsilon) = \mu(|f|^p \geq \varepsilon^p) \leq \frac{1}{\varepsilon^p} \int_X |f|^p \, d\mu = \frac{1}{\varepsilon^p} \|f\|_p^p$$

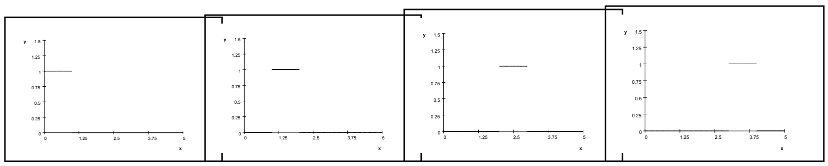
and therefore if $\{f_n\}$ is L^p – Cauchy, then

$$\mu(|f_n - f_m| \geq \varepsilon) \leq \frac{1}{\varepsilon^p} \|f_n - f_m\|_p^p \to 0 \text{ as } m, n \to \infty$$

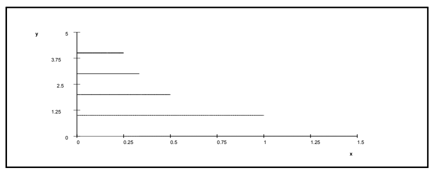
showing $\{f_n\}$ is L^0 – Cauchy. A similar argument holds for the L^p – convergent case. ■



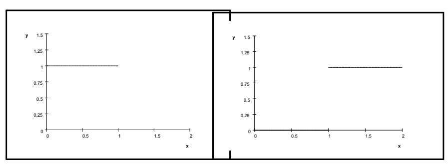
Here is a sequence of functions where $f_n \to 0$ a.e., $f_n \nrightarrow 0$ in L^1, $f_n \overset{m}{\to} 0$.

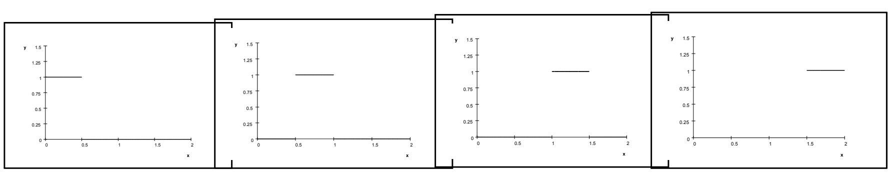


Above is a sequence of functions where $f_n \to 0$ a.e., yet $f_n \nrightarrow 0$ in L^1. or in measure.



Here is a sequence of functions where $f_n \to 0$ a.e., $f_n \overset{m}{\to} 0$ but $f_n \nrightarrow 0$ in L^1.





Above is a sequence of functions where $f_n \to 0$ in L^1, $f_n \nrightarrow 0$ a.e., and $f_n \overset{m}{\to} 0$.

Lemma 21.15. *Suppose $a_n \in \mathbb{C}$ and $|a_{n+1} - a_n| \le \varepsilon_n$ and $\sum_{n=1}^{\infty} \varepsilon_n < \infty$. Then $\lim_{n\to\infty} a_n = a \in \mathbb{C}$ exists and $|a - a_n| \le \delta_n := \sum_{k=n}^{\infty} \varepsilon_k$.*

Proof. (This is a special case of Exercise 6.9.) Let $m > n$ then

$$|a_m - a_n| = \left|\sum_{k=n}^{m-1}(a_{k+1} - a_k)\right| \le \sum_{k=n}^{m-1}|a_{k+1} - a_k| \le \sum_{k=n}^{\infty}\varepsilon_k := \delta_n \,. \quad (21.17)$$

So $|a_m - a_n| \le \delta_{\min(m,n)} \to 0$ as $, m, n \to \infty$, i.e. $\{a_n\}$ is Cauchy. Let $m \to \infty$ in (21.17) to find $|a - a_n| \le \delta_n$. ■

Theorem 21.16. *Suppose $\{f_n\}$ is L^0-Cauchy. Then there exists a subsequence $g_j = f_{n_j}$ of $\{f_n\}$ such that $\lim g_j := f$ exists a.e. and $f_n \xrightarrow{\mu} f$ as $n \to \infty$. Moreover if g is a measurable function such that $f_n \xrightarrow{\mu} g$ as $n \to \infty$, then $f = g$ a.e.*

Proof. Let $\varepsilon_n > 0$ such that $\sum_{n=1}^{\infty}\varepsilon_n < \infty$ ($\varepsilon_n = 2^{-n}$ would do) and set $\delta_n = \sum_{k=n}^{\infty}\varepsilon_k$. Choose $g_j = f_{n_j}$ such that $\{n_j\}$ is a subsequence of $\mathbb{N}$ and

$$\mu(\{|g_{j+1} - g_j| > \varepsilon_j\}) \le \varepsilon_j.$$

Let $E_j = \{|g_{j+1} - g_j| > \varepsilon_j\}$,

$$F_N = \bigcup_{j=N}^{\infty} E_j = \bigcup_{j=N}^{\infty}\{|g_{j+1} - g_j| > \varepsilon_j\}$$

and

$$E := \bigcap_{N=1}^{\infty} F_N = \bigcap_{N=1}^{\infty}\bigcup_{j=N}^{\infty} E_j = \{|g_{j+1} - g_j| > \varepsilon_j \text{ i.o.}\}.$$

Then $\mu(E) = 0$ by Lemma 19.20 or the computation

$$\mu(E) \le \sum_{j=N}^{\infty}\mu(E_j) \le \sum_{j=N}^{\infty}\varepsilon_j = \delta_N \to 0 \text{ as } N \to \infty.$$

If $x \notin F_N$, i.e. $|g_{j+1}(x) - g_j(x)| \le \varepsilon_j$ for all $j \ge N$, then by Lemma 21.15, $f(x) = \lim_{j\to\infty} g_j(x)$ exists and $|f(x) - g_j(x)| \le \delta_j$ for all $j \ge N$. Therefore, since $E^c = \bigcup_{N=1}^{\infty} F_N^c$, $\lim_{j\to\infty} g_j(x) = f(x)$ exists for all $x \notin E$. Moreover, $\{x : |f(x) - g_j(x)| > \delta_j\} \subset F_j$ for all $j \ge N$ and hence

$$\mu(|f - g_j| > \delta_j) \le \mu(F_j) \le \delta_j \to 0 \text{ as } j \to \infty.$$

Therefore $g_j \xrightarrow{\mu} f$ as $j \to \infty$. Since

$$\begin{aligned}\{|f_n - f| > \varepsilon\} &= \{|f - g_j + g_j - f_n| > \varepsilon\} \\ &\subset \{|f - g_j| > \varepsilon/2\} \cup \{|g_j - f_n| > \varepsilon/2\},\end{aligned}$$

$$\mu(\{|f_n - f| > \varepsilon\}) \le \mu(\{|f - g_j| > \varepsilon/2\}) + \mu(|g_j - f_n| > \varepsilon/2)$$

and

$$\mu(\{|f_n - f| > \varepsilon\}) \le \lim_{j\to\infty}\sup \mu(|g_j - f_n| > \varepsilon/2) \to 0 \text{ as } n \to \infty.$$

If there is another function g such that $f_n \xrightarrow{\mu} g$ as $n \to \infty$, then arguing as above

$$\mu(|f - g| > \varepsilon) \le \mu(\{|f - f_n| > \varepsilon/2\}) + \mu(|g - f_n| > \varepsilon/2) \to 0 \text{ as } n \to \infty.$$

Hence

$$\mu(|f - g| > 0) = \mu(\cup_{n=1}^{\infty}\{|f - g| > \frac{1}{n}\}) \le \sum_{n=1}^{\infty}\mu(|f - g| > \frac{1}{n}) = 0,$$

i.e. $f = g$ a.e. ■

Corollary 21.17 (Dominated Convergence Theorem). *Suppose* $\{f_n\}$, $\{g_n\}$, *and* g *are in* L^1 *and* $f \in L^0$ *are functions such that*

$$|f_n| \le g_n \text{ a.e., } f_n \xrightarrow{\mu} f,\ g_n \xrightarrow{\mu} g, \text{ and } \int g_n \to \int g \text{ as } n \to \infty.$$

Then $f \in L^1$ *and* $\lim_{n\to\infty}\|f - f_n\|_1 = 0$, *i.e.* $f_n \to f$ *in* L^1. *In particular* $\lim_{n\to\infty}\int f_n = \int f$.

Proof. First notice that $|f| \le g$ a.e. and hence $f \in L^1$ since $g \in L^1$. To see that $|f| \le g$, use Theorem 21.16 to find subsequences $\{f_{n_k}\}$ and $\{g_{n_k}\}$ of $\{f_n\}$ and $\{g_n\}$ respectively which are almost everywhere convergent. Then

$$|f| = \lim_{k\to\infty}|f_{n_k}| \le \lim_{k\to\infty} g_{n_k} = g \text{ a.e.}$$

If (for sake of contradiction) $\lim_{n\to\infty}\|f - f_n\|_1 \ne 0$ there exists $\varepsilon > 0$ and a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that

$$\int |f - f_{n_k}| \ge \varepsilon \text{ for all } k. \tag{21.18}$$

Using Theorem 21.16 again, we may assume (by passing to a further subsequences if necessary) that $f_{n_k} \to f$ and $g_{n_k} \to g$ almost everywhere. Noting, $|f - f_{n_k}| \le g + g_{n_k} \to 2g$ and $\int (g + g_{n_k}) \to \int 2g$, an application of the dominated convergence Theorem 19.38 implies $\lim_{k\to\infty}\int |f - f_{n_k}| = 0$ which contradicts Eq. (21.18). ■

Exercise 21.1 (Fatou's Lemma). If $f_n \ge 0$ and $f_n \to f$ in measure, then $\int f \le \liminf_{n\to\infty}\int f_n$.

Theorem 21.18 (Egoroff's Theorem). *Suppose $\mu(X) < \infty$ and $f_n \to f$ a.e. Then for all $\varepsilon > 0$ there exists $E \in \mathcal{M}$ such that $\mu(E) < \varepsilon$ and $f_n \to f$ uniformly on E^c. In particular $f_n \xrightarrow{\mu} f$ as $n \to \infty$.*

Proof. Let $f_n \to f$ a.e. Then $\mu(\{|f_n - f| > \frac{1}{k} \text{ i.o. } n\}) = 0$ for all $k > 0$, i.e.

$$\lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} \{|f_n - f| > \frac{1}{k}\}\right) = \mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|f_n - f| > \frac{1}{k}\}\right) = 0.$$

Let $E_k := \bigcup_{n\geq N_k} \{|f_n - f| > \frac{1}{k}\}$ and choose an increasing sequence $\{N_k\}_{k=1}^{\infty}$ such that $\mu(E_k) < \varepsilon 2^{-k}$ for all k. Setting $E := \cup E_k$, $\mu(E) < \sum_k \varepsilon 2^{-k} = \varepsilon$ and if $x \notin E$, then $|f_n - f| \leq \frac{1}{k}$ for all $n \geq N_k$ and all k. That is $f_n \to f$ uniformly on E^c. ∎

Exercise 21.2. Show that Egoroff's Theorem remains valid when the assumption $\mu(X) < \infty$ is replaced by the assumption that $|f_n| \leq g \in L^1$ for all n.
Hint: make use of Theorem 21.18 applied to $f_n|_{X_k}$ where $X_k := \left\{|g| \geq k^{-1}\right\}$.

21.3 Completeness of L^p – spaces

Theorem 21.19. *Let $\|\cdot\|_\infty$ be as defined in Eq. (21.2), then $(L^\infty(X, \mathcal{M}, \mu), \|\cdot\|_\infty)$ is a Banach space. A sequence $\{f_n\}_{n=1}^{\infty} \subset L^\infty$ converges to $f \in L^\infty$ iff there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c. Moreover, bounded simple functions are dense in L^∞.*

Proof. By Minkowski's Theorem 21.4, $\|\cdot\|_\infty$ satisfies the triangle inequality. The reader may easily check the remaining conditions that ensure $\|\cdot\|_\infty$ is a norm. Suppose that $\{f_n\}_{n=1}^{\infty} \subset L^\infty$ is a sequence such $f_n \to f \in L^\infty$, i.e. $\|f - f_n\|_\infty \to 0$ as $n \to \infty$. Then for all $k \in \mathbb{N}$, there exists $N_k < \infty$ such that

$$\mu\left(|f - f_n| > k^{-1}\right) = 0 \text{ for all } n \geq N_k.$$

Let

$$E = \cup_{k=1}^{\infty} \cup_{n\geq N_k} \left\{|f - f_n| > k^{-1}\right\}.$$

Then $\mu(E) = 0$ and for $x \in E^c$, $|f(x) - f_n(x)| \leq k^{-1}$ for all $n \geq N_k$. This shows that $f_n \to f$ uniformly on E^c. Conversely, if there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c, then for any $\varepsilon > 0$,

$$\mu\left(|f - f_n| \geq \varepsilon\right) = \mu\left(\{|f - f_n| \geq \varepsilon\} \cap E^c\right) = 0$$

for all n sufficiently large. That is to say $\limsup_{n\to\infty} \|f - f_n\|_\infty \leq \varepsilon$ for all $\varepsilon > 0$. The density of simple functions follows from the approximation

Theorem 18.42. So the last item to prove is the completeness of L^∞ for which we will use Theorem 7.13.

Suppose that $\{f_n\}_{n=1}^\infty \subset L^\infty$ is a sequence such that $\sum_{n=1}^\infty \|f_n\|_\infty < \infty$. Let $M_n := \|f_n\|_\infty$, $E_n := \{|f_n| > M_n\}$, and $E := \cup_{n=1}^\infty E_n$ so that $\mu(E) = 0$. Then

$$\sum_{n=1}^\infty \sup_{x\in E^c} |f_n(x)| \le \sum_{n=1}^\infty M_n < \infty$$

which shows that $S_N(x) = \sum_{n=1}^N f_n(x)$ converges uniformly to $S(x) := \sum_{n=1}^\infty f_n(x)$ on E^c, i.e. $\lim_{n\to\infty} \|S - S_n\|_\infty = 0$.

Alternatively, suppose $\varepsilon_{m,n} := \|f_m - f_n\|_\infty \to 0$ as $m, n \to \infty$. Let $E_{m,n} = \{|f_n - f_m| > \varepsilon_{m,n}\}$ and $E := \cup E_{m,n}$, then $\mu(E) = 0$ and

$$\sup_{x\in E^c} |f_m(x) - f_n(x)| \le \varepsilon_{m,n} \to 0 \text{ as } m, n \to \infty.$$

Therefore, $f := \lim_{n\to\infty} f_n$ exists on E^c and the limit is uniform on E^c. Letting $f = \lim_{n\to\infty} 1_{E^c} f_n$, it then follows that $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$. ■

Theorem 21.20 (Completeness of $L^p(\mu)$). *For $1 \le p \le \infty$, $L^p(\mu)$ equipped with the L^p – norm, $\|\cdot\|_p$ (see Eq. (21.1)), is a Banach space.*

Proof. By Minkowski's Theorem 21.4, $\|\cdot\|_p$ satisfies the triangle inequality. As above the reader may easily check the remaining conditions that ensure $\|\cdot\|_p$ is a norm. So we are left to prove the completeness of $L^p(\mu)$ for $1 \le p < \infty$, the case $p = \infty$ being done in Theorem 21.19.

Let $\{f_n\}_{n=1}^\infty \subset L^p(\mu)$ be a Cauchy sequence. By Chebyshev's inequality (Lemma 21.14), $\{f_n\}$ is L^0-Cauchy (i.e. Cauchy in measure) and by Theorem 21.16 there exists a subsequence $\{g_j\}$ of $\{f_n\}$ such that $g_j \to f$ a.e. By Fatou's Lemma,

$$\begin{aligned}\|g_j - f\|_p^p &= \int \lim_{k\to\infty} \inf |g_j - g_k|^p d\mu \le \lim_{k\to\infty} \inf \int |g_j - g_k|^p d\mu \\ &= \lim_{k\to\infty} \inf \|g_j - g_k\|_p^p \to 0 \text{ as } j \to \infty.\end{aligned}$$

In particular, $\|f\|_p \le \|g_j - f\|_p + \|g_j\|_p < \infty$ so the $f \in L^p$ and $g_j \xrightarrow{L^p} f$. The proof is finished because,

$$\|f_n - f\|_p \le \|f_n - g_j\|_p + \|g_j - f\|_p \to 0 \text{ as } j, n \to \infty.$$

■

The $L^p(\mu)$ – norm controls two types of behaviors of f, namely the "behavior at infinity" and the behavior of "local singularities." So in particular, if f is blows up at a point $x_0 \in X$, then locally near x_0 it is harder for f to be in $L^p(\mu)$ as p increases. On the other hand a function $f \in L^p(\mu)$ is allowed to decay at "infinity" slower and slower as p increases. With these insights in mind, we should not in general expect $L^p(\mu) \subset L^q(\mu)$ or $L^q(\mu) \subset L^p(\mu)$. However, there are two notable exceptions. (1) If $\mu(X) < \infty$, then there is no behavior at infinity to worry about and $L^q(\mu) \subset L^p(\mu)$ for all $q \le p$ as is shown in Corollary 21.21 below. (2) If μ is counting measure, i.e. $\mu(A) = \#(A)$, then all functions in $L^p(\mu)$ for any p can not blow up on a set of positive measure, so there are no local singularities. In this case $L^p(\mu) \subset L^q(\mu)$ for all $q \le p$, see Corollary 21.25 below.

Corollary 21.21. *If $\mu(X) < \infty$ and $0 < p < q \le \infty$, then $L^q(\mu) \subset L^p(\mu)$, the inclusion map is bounded and in fact*

$$\|f\|_p \le [\mu(X)]^{\left(\frac{1}{p}-\frac{1}{q}\right)} \|f\|_q .$$

Proof. Take $a \in [1, \infty]$ such that

$$\frac{1}{p} = \frac{1}{a} + \frac{1}{q}, \text{ i.e. } a = \frac{pq}{q-p}.$$

Then by Corollary 21.3,

$$\|f\|_p = \|f \cdot 1\|_p \le \|f\|_q \cdot \|1\|_a = \mu(X)^{1/a} \|f\|_q = \mu(X)^{\left(\frac{1}{p}-\frac{1}{q}\right)} \|f\|_q .$$

The reader may easily check this final formula is correct even when $q = \infty$ provided we interpret $1/p - 1/\infty$ to be $1/p$. ■

Proposition 21.22. *Suppose that $0 < p_0 < p_1 \le \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ be defined by*

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1} \tag{21.19}$$

with the interpretation that $\lambda/p_1 = 0$ if $p_1 = \infty$.[1] *Then $L^{p_\lambda} \subset L^{p_0} + L^{p_1}$, i.e. every function $f \in L^{p_\lambda}$ may be written as $f = g + h$ with $g \in L^{p_0}$ and $h \in L^{p_1}$. For $1 \le p_0 < p_1 \le \infty$ and $f \in L^{p_0} + L^{p_1}$ let*

$$\|f\| := \inf \left\{ \|g\|_{p_0} + \|h\|_{p_1} : f = g + h \right\}.$$

Then $(L^{p_0} + L^{p_1}, \|\cdot\|)$ is a Banach space and the inclusion map from L^{p_λ} to $L^{p_0} + L^{p_1}$ is bounded; in fact $\|f\| \le 2 \|f\|_{p_\lambda}$ for all $f \in L^{p_\lambda}$.

Proof. Let $M > 0$, then the local singularities of f are contained in the set $E := \{|f| > M\}$ and the behavior of f at "infinity" is solely determined by f on E^c. Hence let $g = f 1_E$ and $h = f 1_{E^c}$ so that $f = g + h$. By our earlier discussion we expect that $g \in L^{p_0}$ and $h \in L^{p_1}$ and this is the case since,

[1] A little algebra shows that λ may be computed in terms of p_0, p_λ and p_1 by

$$\lambda = \frac{p_0}{p_\lambda} \cdot \frac{p_1 - p_\lambda}{p_1 - p_0}.$$

$$\|g\|_{p_0}^{p_0} = \int |f|^{p_0} 1_{|f|>M} = M^{p_0} \int \left|\frac{f}{M}\right|^{p_0} 1_{|f|>M}$$
$$\leq M^{p_0} \int \left|\frac{f}{M}\right|^{p_\lambda} 1_{|f|>M} \leq M^{p_0-p_\lambda} \|f\|_{p_\lambda}^{p_\lambda} < \infty$$

and

$$\|h\|_{p_1}^{p_1} = \left\|f 1_{|f|\leq M}\right\|_{p_1}^{p_1} = \int |f|^{p_1} 1_{|f|\leq M} = M^{p_1} \int \left|\frac{f}{M}\right|^{p_1} 1_{|f|\leq M}$$
$$\leq M^{p_1} \int \left|\frac{f}{M}\right|^{p_\lambda} 1_{|f|\leq M} \leq M^{p_1-p_\lambda} \|f\|_{p_\lambda}^{p_\lambda} < \infty.$$

Moreover this shows

$$\|f\| \leq M^{1-p_\lambda/p_0} \|f\|_{p_\lambda}^{p_\lambda/p_0} + M^{1-p_\lambda/p_1} \|f\|_{p_\lambda}^{p_\lambda/p_1}.$$

Taking $M = \lambda \|f\|_{p_\lambda}$ then gives

$$\|f\| \leq \left(\lambda^{1-p_\lambda/p_0} + \lambda^{1-p_\lambda/p_1}\right) \|f\|_{p_\lambda}$$

and then taking $\lambda = 1$ shows $\|f\| \leq 2\|f\|_{p_\lambda}$. The the proof that $(L^{p_0} + L^{p_1}, \|\cdot\|)$ is a Banach space is left as Exercise 21.7 to the reader. ■

Corollary 21.23 (Interpolation of L^p – norms). *Suppose that $0 < p_0 < p_1 \leq \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ be defined as in Eq. (21.19), then $L^{p_0} \cap L^{p_1} \subset L^{p_\lambda}$ and*

$$\|f\|_{p_\lambda} \leq \|f\|_{p_0}^{\lambda} \|f\|_{p_1}^{1-\lambda}. \tag{21.20}$$

Further assume $1 \leq p_0 < p_\lambda < p_1 \leq \infty$, and for $f \in L^{p_0} \cap L^{p_1}$ let

$$\|f\| := \|f\|_{p_0} + \|f\|_{p_1}.$$

Then $(L^{p_0} \cap L^{p_1}, \|\cdot\|)$ is a Banach space and the inclusion map of $L^{p_0} \cap L^{p_1}$ into L^{p_λ} is bounded, in fact

$$\|f\|_{p_\lambda} \leq \max\left(\lambda^{-1}, (1-\lambda)^{-1}\right) \left(\|f\|_{p_0} + \|f\|_{p_1}\right). \tag{21.21}$$

The heuristic explanation of this corollary is that if $f \in L^{p_0} \cap L^{p_1}$, then f has local singularities no worse than an L^{p_1} function and behavior at infinity no worse than an L^{p_0} function. Hence $f \in L^{p_\lambda}$ for any p_λ between p_0 and p_1.

Proof. Let λ be determined as above, $a = p_0/\lambda$ and $b = p_1/(1-\lambda)$, then by Corollary 21.3,

$$\|f\|_{p_\lambda} = \left\||f|^\lambda |f|^{1-\lambda}\right\|_{p_\lambda} \leq \left\||f|^\lambda\right\|_a \left\||f|^{1-\lambda}\right\|_b = \|f\|_{p_0}^\lambda \|f\|_{p_1}^{1-\lambda}.$$

It is easily checked that $\|\cdot\|$ is a norm on $L^{p_0} \cap L^{p_1}$. To show this space is complete, suppose that $\{f_n\} \subset L^{p_0} \cap L^{p_1}$ is a $\|\cdot\|$ – Cauchy sequence. Then

$\{f_n\}$ is both L^{p_0} and L^{p_1} – Cauchy. Hence there exist $f \in L^{p_0}$ and $g \in L^{p_1}$ such that $\lim_{n\to\infty} \|f - f_n\|_{p_0} = 0$ and $\lim_{n\to\infty} \|g - f_n\|_{p_\lambda} = 0$. By Chebyshev's inequality (Lemma 21.14) $f_n \to f$ and $f_n \to g$ in measure and therefore by Theorem 21.16, $f = g$ a.e. It now is clear that $\lim_{n\to\infty} \|f - f_n\| = 0$. The estimate in Eq. (21.21) is left as Exercise 21.6 to the reader. ∎

Remark 21.24. Combining Proposition 21.22 and Corollary 21.23 gives

$$L^{p_0} \cap L^{p_1} \subset L^{p_\lambda} \subset L^{p_0} + L^{p_1}$$

for $0 < p_0 < p_1 \le \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ as in Eq. (21.19).

Corollary 21.25. *Suppose now that μ is counting measure on X. Then $L^p(\mu) \subset L^q(\mu)$ for all $0 < p < q \le \infty$ and $\|f\|_q \le \|f\|_p$.*

Proof. Suppose that $0 < p < q = \infty$, then

$$\|f\|_\infty^p = \sup\{|f(x)|^p : x \in X\} \le \sum_{x\in X} |f(x)|^p = \|f\|_p^p,$$

i.e. $\|f\|_\infty \le \|f\|_p$ for all $0 < p < \infty$. For $0 < p \le q \le \infty$, apply Corollary 21.23 with $p_0 = p$ and $p_1 = \infty$ to find

$$\|f\|_q \le \|f\|_p^{p/q} \|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q} \|f\|_p^{1-p/q} = \|f\|_p.$$

∎

21.3.1 Summary:

1. Since $\mu(|f| > \varepsilon) \le \varepsilon^{-p} \|f\|_p^p$, L^p – convergence implies L^0 – convergence.
2. L^0 – convergence implies almost everywhere convergence for some subsequence.
3. If $\mu(X) < \infty$ then almost everywhere convergence implies uniform convergence off certain sets of small measure and in particular we have L^0 – convergence.
4. If $\mu(X) < \infty$, then $L^q \subset L^p$ for all $p \le q$ and L^q – convergence implies L^p – convergence.
5. $L^{p_0} \cap L^{p_1} \subset L^q \subset L^{p_0} + L^{p_1}$ for any $q \in (p_0, p_1)$.
6. If $p \le q$, then $\ell^p \subset \ell^q$ and $\|f\|_q \le \|f\|_p$.

21.4 Converse of Hölder's Inequality

Throughout this section we assume $(X, \mathcal{M}, \mu)$ is a σ – finite measure space, $q \in [1,\infty]$ and $p \in [1,\infty]$ are conjugate exponents, i.e. $p^{-1} + q^{-1} = 1$. For $g \in L^q$, let $\phi_g \in (L^p)^*$ be given by

$$\phi_g(f) = \int gf \; d\mu. \tag{21.22}$$

By Hölder's inequality

$$|\phi_g(f)| \le \int |gf| d\mu \le \|g\|_q \|f\|_p \tag{21.23}$$

which implies that

$$\|\phi_g\|_{(L^p)^*} := \sup\{|\phi_g(f)| : \|f\|_p = 1\} \le \|g\|_q. \tag{21.24}$$

Proposition 21.26 (Converse of Hölder's Inequality). *Let $(X, \mathcal{M}, \mu)$ be a σ – finite measure space and $1 \le p \le \infty$ as above. For all $g \in L^q$,*

$$\|g\|_q = \|\phi_g\|_{(L^p)^*} := \sup\left\{|\phi_g(f)| : \|f\|_p = 1\right\} \tag{21.25}$$

and for any measurable function $g : X \to \mathbb{C}$,

$$\|g\|_q = \sup\left\{\int_X |g| \, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\}. \tag{21.26}$$

Proof. We begin by proving Eq. (21.25). Assume first that $q < \infty$ so $p > 1$. Then

$$|\phi_g(f)| = \left|\int gf \; d\mu\right| \le \int |gf| \, d\mu \le \|g\|_q \|f\|_p$$

and equality occurs in the first inequality when $\mathrm{sgn}(gf)$ is constant a.e. while equality in the second occurs, by Theorem 21.2, when $|f|^p = c|g|^q$ for some constant $c > 0$. So let $f := \overline{\mathrm{sgn}(g)}|g|^{q/p}$ which for $p = \infty$ is to be interpreted as $f = \overline{\mathrm{sgn}(g)}$, i.e. $|g|^{q/\infty} \equiv 1$. When $p = \infty$,

$$|\phi_g(f)| = \int_X g \, \overline{\mathrm{sgn}(g)} d\mu = \|g\|_{L^1(\mu)} = \|g\|_1 \, \|f\|_\infty$$

which shows that $\|\phi_g\|_{(L^\infty)^*} \ge \|g\|_1$. If $p < \infty$, then

$$\|f\|_p^p = \int |f|^p = \int |g|^q = \|g\|_q^q$$

while

$$\phi_g(f) = \int gf d\mu = \int |g||g|^{q/p} d\mu = \int |g|^q d\mu = \|g\|_q^q.$$

Hence

$$\frac{|\phi_g(f)|}{\|f\|_p} = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q^{q(1-\frac{1}{p})} = \|g\|_q.$$

This shows that $||\phi_g|| \ge \|g\|_q$ which combined with Eq. (21.24) implies Eq. (21.25).

The last case to consider is $p = 1$ and $q = \infty$. Let $M := \|g\|_\infty$ and choose $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ as $n \to \infty$ and $\mu(X_n) < \infty$ for all n. For any $\varepsilon > 0$, $\mu(|g| \geq M - \varepsilon) > 0$ and $X_n \cap \{|g| \geq M - \varepsilon\} \uparrow \{|g| \geq M - \varepsilon\}$. Therefore, $\mu(X_n \cap \{|g| \geq M - \varepsilon\}) > 0$ for n sufficiently large. Let

$$f = \overline{\operatorname{sgn}(g)} 1_{X_n \cap \{|g| \geq M - \varepsilon\}},$$

then

$$\|f\|_1 = \mu(X_n \cap \{|g| \geq M - \varepsilon\}) \in (0, \infty)$$

and

$$\begin{aligned} |\phi_g(f)| &= \int_{X_n \cap \{|g| \geq M - \varepsilon\}} \overline{\operatorname{sgn}(g)} g d\mu = \int_{X_n \cap \{|g| \geq M - \varepsilon\}} |g| d\mu \\ &\geq (M - \varepsilon)\mu(X_n \cap \{|g| \geq M - \varepsilon\}) = (M - \varepsilon)\|f\|_1. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows from this equation that $\|\phi_g\|_{(L^1)^*} \geq M = \|g\|_\infty$.

Now for the proof of Eq. (21.26). The key new point is that we no longer are assuming that $g \in L^q$. Let $M(g)$ denote the right member in Eq. (21.26) and set $g_n := 1_{X_n \cap \{|g| \leq n\}} g$. Then $|g_n| \uparrow |g|$ as $n \to \infty$ and it is clear that $M(g_n)$ is increasing in n. Therefore using Lemma 4.10 and the monotone convergence theorem,

$$\begin{aligned} \lim_{n \to \infty} M(g_n) &= \sup_n M(g_n) = \sup_n \sup \left\{ \int_X |g_n| f d\mu : \|f\|_p = 1 \text{ and } f \geq 0 \right\} \\ &= \sup \left\{ \sup_n \int_X |g_n| f d\mu : \|f\|_p = 1 \text{ and } f \geq 0 \right\} \\ &= \sup \left\{ \lim_{n \to \infty} \int_X |g_n| f d\mu : \|f\|_p = 1 \text{ and } f \geq 0 \right\} \\ &= \sup \left\{ \int_X |g| f d\mu : \|f\|_p = 1 \text{ and } f \geq 0 \right\} = M(g). \end{aligned}$$

Since $g_n \in L^q$ for all n and $M(g_n) = \|\phi_{g_n}\|_{(L^p)^*}$ (as you should verify), it follows from Eq. (21.25) that $M(g_n) = \|g_n\|_q$. When $q < \infty$ (by the monotone convergence theorem) and when $q = \infty$ (directly from the definitions) one learns that $\lim_{n \to \infty} \|g_n\|_q = \|g\|_q$. Combining this fact with $\lim_{n \to \infty} M(g_n) = M(g)$ just proved shows $M(g) = \|g\|_q$. ■

As an application we can derive a sweeping generalization of Minkowski's inequality. (See Reed and Simon, Vol II. Appendix IX.4 for a more thorough discussion of complex interpolation theory.)

Theorem 21.27 (Minkowski's Inequality for Integrals). *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ – finite measure spaces and $1 \leq p \leq \infty$. If f is a $\mathcal{M} \otimes \mathcal{N}$ measurable function, then $y \to \|f(\cdot, y)\|_{L^p(\mu)}$ is measurable and*

1. if f is a positive $\mathcal{M}\otimes\mathcal{N}$ measurable function, then

$$\|\int_Y f(\cdot,y)d\nu(y)\|_{L^p(\mu)} \le \int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y). \tag{21.27}$$

2. If $f : X\times Y \to \mathbb{C}$ is a $\mathcal{M}\otimes\mathcal{N}$ measurable function and $\int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y) < \infty$ then
a) for μ – a.e. x, $f(x,\cdot) \in L^1(\nu)$,
b) the μ –a.e. defined function, $x \to \int_Y f(x,y)d\nu(y)$, is in $L^p(\mu)$ and
c) the bound in Eq. (21.27) holds.

Proof. For $p \in [1,\infty]$, let $F_p(y) := \|f(\cdot,y)\|_{L^p(\mu)}$. If $p \in [1,\infty)$

$$F_p(y) = \|f(\cdot,y)\|_{L^p(\mu)} = \left(\int_X |f(x,y)|^p \, d\mu(x)\right)^{1/p}$$

is a measurable function on Y by Fubini's theorem. To see that F_∞ is measurable, let $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ for all n. Then by Exercise 21.5,

$$F_\infty(y) = \lim_{n\to\infty}\lim_{p\to\infty}\|f(\cdot,y)1_{X_n}\|_{L^p(\mu)}$$

which shows that F_∞ is $(Y,\mathcal{N})$ – measurable as well. This shows that integral on the right side of Eq. (21.27) is well defined.

Now suppose that $f \ge 0$, $q = p/(p-1)$and $g \in L^q(\mu)$ such that $g \ge 0$ and $\|g\|_{L^q(\mu)} = 1$. Then by Tonelli's theorem and Hölder's inequality,

$$\begin{aligned}\int_X \left[\int_Y f(x,y)d\nu(y)\right] g(x)d\mu(x) &= \int_Y d\nu(y)\int_X d\mu(x)f(x,y)g(x) \\ &\le \|g\|_{L^q(\mu)}\int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y) \\ &= \int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y).\end{aligned}$$

Therefore by the converse to Hölder's inequality (Proposition 21.26),

$$\begin{aligned}&\|\int_Y f(\cdot,y)d\nu(y)\|_{L^p(\mu)} \\ &= \sup\left\{\int_X\left[\int_Y f(x,y)d\nu(y)\right]g(x)d\mu(x) : \|g\|_{L^q(\mu)} = 1 \text{ and } g \ge 0\right\} \\ &\le \int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y)\end{aligned}$$

proving Eq. (21.27) in this case.

Now let $f : X\times Y \to \mathbb{C}$ be as in item 2) of the theorem. Applying the first part of the theorem to $|f|$ shows

$$\int_Y |f(x,y)|\, d\nu(y) < \infty \text{ for } \mu\text{– a.e. } x,$$

i.e. $f(x,\cdot) \in L^1(\nu)$ for the μ –a.e. x. Since $\left|\int_Y f(x,y)d\nu(y)\right| \le \int_Y |f(x,y)|\, d\nu(y)$ it follows by item 1) that

$$\left\|\int_Y f(\cdot,y)d\nu(y)\right\|_{L^p(\mu)} \le \left\|\int_Y |f(\cdot,y)|\, d\nu(y)\right\|_{L^p(\mu)} \le \int_Y \|f(\cdot,y)\|_{L^p(\mu)}\, d\nu(y).$$

Hence the function, $x \in X \to \int_Y f(x,y)d\nu(y)$, is in $L^p(\mu)$ and the bound in Eq. (21.27) holds. ■

Here is an application of Minkowski's inequality for integrals. In this theorem we will be using the convention that $x^{-1/\infty} := 1$.

Theorem 21.28 (Theorem 6.20 in Folland). *Suppose that $k : (0,\infty) \times (0,\infty) \to \mathbb{C}$ is a measurable function such that k is homogenous of degree -1, i.e. $k(\lambda x, \lambda y) = \lambda^{-1} k(x,y)$ for all $\lambda > 0$. If, for some $p \in [1,\infty]$,*

$$C_p := \int_0^\infty |k(x,1)|\, x^{-1/p} dx < \infty$$

then for $f \in L^p((0,\infty),m)$, $k(x,\cdot)f(\cdot) \in L^1((0,\infty),m)$ for m – a.e. x. Moreover, the m – a.e. defined function

$$(Kf)(x) = \int_0^\infty k(x,y)f(y)dy \tag{21.28}$$

is in $L^p((0,\infty),m)$ and

$$\|Kf\|_{L^p((0,\infty),m)} \le C_p \|f\|_{L^p((0,\infty),m)}.$$

Proof. By the homogeneity of k, $k(x,y) = x^{-1}k(1,\frac{y}{x})$. Using this relation and making the change of variables, $y = zx$, gives

$$\begin{aligned}\int_0^\infty |k(x,y)f(y)|\, dy &= \int_0^\infty x^{-1}\left|k(1,\frac{y}{x})f(y)\right| dy \\ &= \int_0^\infty x^{-1}\,|k(1,z)f(xz)|\, xdz = \int_0^\infty |k(1,z)f(xz)|\, dz.\end{aligned}$$

Since

$$\|f(\cdot\ z)\|^p_{L^p((0,\infty),m)} = \int_0^\infty |f(yz)|^p dy = \int_0^\infty |f(x)|^p\ \frac{dx}{z},$$

$$\|f(\cdot\ z)\|_{L^p((0,\infty),m)} = z^{-1/p}\|f\|_{L^p((0,\infty),m)}.$$

Using Minkowski's inequality for integrals then shows

$$
\begin{aligned}
\left\| \int_0^\infty |k(\cdot, y) f(y)|\, dy \right\|_{L^p((0,\infty),m)} &\leq \int_0^\infty |k(1,z)|\, \|f(\cdot z)\|_{L^p((0,\infty),m)}\, dz \\
&= \|f\|_{L^p((0,\infty),m)} \int_0^\infty |k(1,z)|\, z^{-1/p} dz \\
&= C_p \|f\|_{L^p((0,\infty),m)} < \infty.
\end{aligned}
$$

This shows that Kf in Eq. (21.28) is well defined from m – a.e. x. The proof is finished by observing

$$
\|Kf\|_{L^p((0,\infty),m)} \leq \left\| \int_0^\infty |k(\cdot, y) f(y)|\, dy \right\|_{L^p((0,\infty),m)} \leq C_p \|f\|_{L^p((0,\infty),m)}
$$

for all $f \in L^p((0,\infty), m)$. ■

The following theorem is a strengthening of Proposition 21.26. It may be skipped on the first reading.

Theorem 21.29 (Converse of Hölder's Inequality II). *Assume that $(X, \mathcal{M}, \mu)$ is a σ – finite measure space, $q, p \in [1, \infty]$ are conjugate exponents and let $\mathbb{S}_f$ denote the set of simple functions ϕ on X such that $\mu(\phi \neq 0) < \infty$. Let $g : X \to \mathbb{C}$ be a measurable function such that $\phi g \in L^1(\mu)$ for all $\phi \in \mathbb{S}_f$,*[2] *and define*

$$
M_q(g) := \sup \left\{ \left| \int_X \phi g d\mu \right| : \phi \in \mathbb{S}_f \text{ with } \|\phi\|_p = 1 \right\}. \tag{21.29}
$$

If $M_q(g) < \infty$ then $g \in L^q(\mu)$ and $M_q(g) = \|g\|_q$.

Proof. Let $X_n \in \mathcal{M}$ be sets such that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \uparrow \infty$. Suppose that $q = 1$ and hence $p = \infty$. Choose simple functions ϕ_n on X such that $|\phi_n| \leq 1$ and $\overline{\operatorname{sgn}(g)} = \lim_{n\to\infty} \phi_n$ in the pointwise sense. Then $1_{X_m} \phi_n \in \mathbb{S}_f$ and therefore

$$
\left| \int_X 1_{X_m} \phi_n g d\mu \right| \leq M_q(g)
$$

for all m, n. By assumption $1_{X_m} g \in L^1(\mu)$ and therefore by the dominated convergence theorem we may let $n \to \infty$ in this equation to find

$$
\int_X 1_{X_m} |g|\, d\mu \leq M_q(g)
$$

for all m. The monotone convergence theorem then implies that

$$
\int_X |g|\, d\mu = \lim_{m\to\infty} \int_X 1_{X_m} |g|\, d\mu \leq M_q(g)
$$

[2] This is equivalent to requiring $1_A g \in L^1(\mu)$ for all $A \in \mathcal{M}$ such that $\mu(A) < \infty$.

showing $g \in L^1(\mu)$ and $\|g\|_1 \le M_q(g)$. Since Holder's inequality implies that $M_q(g) \le \|g\|_1$, we have proved the theorem in case $q = 1$. For $q > 1$, we will begin by assuming that $g \in L^q(\mu)$. Since $p \in [1, \infty)$ we know that $\mathbb{S}_f$ is a dense subspace of $L^p(\mu)$ and therefore, using ϕ_g is continuous on $L^p(\mu)$,

$$M_q(g) = \sup\left\{\left|\int_X \phi g d\mu\right| : \phi \in L^p(\mu) \text{ with } \|\phi\|_p = 1\right\} = \|g\|_q$$

where the last equality follows by Proposition 21.26. So it remains to show that if $\phi g \in L^1$ for all $\phi \in \mathbb{S}_f$ and $M_q(g) < \infty$ then $g \in L^q(\mu)$. For $n \in \mathbb{N}$, let $g_n := 1_{X_n} 1_{|g|\le n} g$. Then $g_n \in L^q(\mu)$, in fact $\|g_n\|_q \le n\mu(X_n)^{1/q} < \infty$. So by the previous paragraph, $\|g_n\|_q = M_q(g_n)$ and hence

$$\begin{aligned}\|g_n\|_q &= \sup\left\{\left|\int_X \phi 1_{X_n} 1_{|g|\le n} g d\mu\right| : \phi \in L^p(\mu) \text{ with } \|\phi\|_p = 1\right\} \\ &\le M_q(g) \left\|\phi 1_{X_n} 1_{|g|\le n}\right\|_p \le M_q(g) \cdot 1 = M_q(g)\end{aligned}$$

wherein the second to last inequality we have made use of the definition of $M_q(g)$ and the fact that $\phi 1_{X_n} 1_{|g|\le n} \in \mathbb{S}_f$. If $q \in (1, \infty)$, an application of the monotone convergence theorem (or Fatou's Lemma) along with the continuity of the norm, $\|\cdot\|_p$, implies

$$\|g\|_q = \lim_{n\to\infty} \|g_n\|_q \le M_q(g) < \infty.$$

If $q = \infty$, then $\|g_n\|_\infty \le M_q(g) < \infty$ for all n implies $|g_n| \le M_q(g)$ a.e. which then implies that $|g| \le M_q(g)$ a.e. since $|g| = \lim_{n\to\infty} |g_n|$. That is $g \in L^\infty(\mu)$ and $\|g\|_\infty \le M_\infty(g)$. ∎

21.5 Uniform Integrability

This section will address the question as to what extra conditions are needed in order that an L^0 – convergent sequence is L^p – convergent.

Notation 21.30 *For $f \in L^1(\mu)$ and $E \in \mathcal{M}$, let*

$$\mu(f : E) := \int_E f d\mu.$$

and more generally if $A, B \in \mathcal{M}$ let

$$\mu(f : A, B) := \int_{A\cap B} f d\mu.$$

Lemma 21.31. *Suppose $g \in L^1(\mu)$, then for any $\varepsilon > 0$ there exist a $\delta > 0$ such that $\mu(|g| : E) < \varepsilon$ whenever $\mu(E) < \delta$.*

Proof. If the Lemma is false, there would exist $\varepsilon > 0$ and sets E_n such that $\mu(E_n) \to 0$ while $\mu(|g| : E_n) \geq \varepsilon$ for all n. Since $|1_{E_n} g| \leq |g| \in L^1$ and for any $\delta \in (0,1)$, $\mu(1_{E_n} |g| > \delta) \leq \mu(E_n) \to 0$ as $n \to \infty$, the dominated convergence theorem of Corollary 21.17 implies $\lim_{n\to\infty} \mu(|g| : E_n) = 0$. This contradicts $\mu(|g| : E_n) \geq \varepsilon$ for all n and the proof is complete. ∎

Suppose that $\{f_n\}_{n=1}^\infty$ is a sequence of measurable functions which converge in $L^1(\mu)$ to a function f. Then for $E \in \mathcal{M}$ and $n \in \mathbb{N}$,

$$|\mu(f_n : E)| \leq |\mu(f - f_n : E)| + |\mu(f : E)| \leq \|f - f_n\|_1 + |\mu(f : E)| .$$

Let $\varepsilon_N := \sup_{n>N} \|f - f_n\|_1$, then $\varepsilon_N \downarrow 0$ as $N \uparrow \infty$ and

$$\sup_n |\mu(f_n : E)| \leq \sup_{n \leq N} |\mu(f_n : E)| \vee (\varepsilon_N + |\mu(f : E)|) \leq \varepsilon_N + \mu\left(g_N : E\right), \tag{21.30}$$

where $g_N = |f| + \sum_{n=1}^N |f_n| \in L^1$. From Lemma 21.31 and Eq. (21.30) one easily concludes,

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \ni\ \sup_n |\mu(f_n : E)| < \varepsilon \text{ when } \mu(E) < \delta. \tag{21.31}$$

Definition 21.32. *Functions $\{f_n\}_{n=1}^\infty \subset L^1(\mu)$ satisfying Eq. (21.31) are said to be* uniformly integrable.

Remark 21.33. Let $\{f_n\}$ be real functions satisfying Eq. (21.31), E be a set where $\mu(E) < \delta$ and $E_n = E \cap \{f_n \geq 0\}$. Then $\mu(E_n) < \delta$ so that $\mu(f_n^+ : E) = \mu(f_n : E_n) < \varepsilon$ and similarly $\mu(f_n^- : E) < \varepsilon$. Therefore if Eq. (21.31) holds then

$$\sup_n \mu(|f_n| : E) < 2\varepsilon \text{ when } \mu(E) < \delta. \tag{21.32}$$

Similar arguments work for the complex case by looking at the real and imaginary parts of f_n. Therefore $\{f_n\}_{n=1}^\infty \subset L^1(\mu)$ is uniformly integrable iff

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \ni\ \sup_n \mu(|f_n| : E) < \varepsilon \text{ when } \mu(E) < \delta. \tag{21.33}$$

Lemma 21.34. *Assume that $\mu(X) < \infty$, then $\{f_n\}$ is uniformly bounded in $L^1(\mu)$ (i.e. $K = \sup_n \|f_n\|_1 < \infty$) and $\{f_n\}$ is uniformly integrable iff*

$$\lim_{M\to\infty} \sup_n \mu(|f_n| : |f_n| \geq M) = 0. \tag{21.34}$$

Proof. Since $\{f_n\}$ is uniformly bounded in $L^1(\mu)$, $\mu(|f_n| \geq M) \leq K/M$. So if (21.33) holds and $\varepsilon > 0$ is given, we may choose M sufficiently large so that $\mu(|f_n| \geq M) < \delta(\varepsilon)$ for all n and therefore,

$$\sup_n \mu(|f_n| : |f_n| \geq M) \leq \varepsilon.$$

Since ε is arbitrary, we concluded that Eq. (21.34) must hold. Conversely, suppose that Eq. (21.34) holds, then automatically $K = \sup_n \mu(|f_n|) < \infty$ because

$$\mu(|f_n|) = \mu(|f_n| : |f_n| \geq M) + \mu(|f_n| : |f_n| < M)$$
$$\leq \sup_n \mu(|f_n| : |f_n| \geq M) + M\mu(X) < \infty.$$

Moreover,

$$\mu(|f_n| : E) = \mu(|f_n| : |f_n| \geq M, E) + \mu(|f_n| : |f_n| < M, E)$$
$$\leq \sup_n \mu(|f_n| : |f_n| \geq M) + M\mu(E).$$

So given $\varepsilon > 0$ choose M so large that $\sup_n \mu(|f_n| : |f_n| \geq M) < \varepsilon/2$ and then take $\delta = \varepsilon / (2M)$. ■

Remark 21.35. It is not in general true that if $\{f_n\} \subset L^1(\mu)$ is uniformly integrable then $\sup_n \mu(|f_n|) < \infty$. For example take $X = \{*\}$ and $\mu(\{*\}) = 1$. Let $f_n(*) = n$. Since for $\delta < 1$ a set $E \subset X$ such that $\mu(E) < \delta$ is in fact the empty set, we see that Eq. (21.32) holds in this example. However, for finite measure spaces with out "atoms", for every $\delta > 0$ we may find a finite partition of X by sets $\{E_\ell\}_{\ell=1}^k$ with $\mu(E_\ell) < \delta$. Then if Eq. (21.32) holds with $2\varepsilon = 1$, then

$$\mu(|f_n|) = \sum_{\ell=1}^k \mu(|f_n| : E_\ell) \leq k$$

showing that $\mu(|f_n|) \leq k$ for all n.

The following Lemmas gives a concrete necessary and sufficient conditions for verifying a sequence of functions is uniformly bounded and uniformly integrable.

Lemma 21.36. *Suppose that $\mu(X) < \infty$, and $\Lambda \subset L^0(X)$ is a collection of functions.*

1. *If there exists a non decreasing function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{x\to\infty} \phi(x)/x = \infty$ and*
$$K := \sup_{f\in\Lambda} \mu(\phi(|f|)) < \infty \tag{21.35}$$
then
$$\lim_{M\to\infty} \sup_{f\in\Lambda} \mu\left(|f| \, 1_{|f|\geq M}\right) = 0. \tag{21.36}$$
2. *Conversely if Eq. (21.36) holds, there exists a non-decreasing continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\phi(0) = 0$, $\lim_{x\to\infty} \phi(x)/x = \infty$ and Eq. (21.35) is valid.*

Proof. 1. Let ϕ be as in item 1. above and set $\varepsilon_M := \sup_{x\geq M} \frac{x}{\phi(x)} \to 0$ as $M \to \infty$ by assumption. Then for $f \in \Lambda$

$$\mu(|f|:|f|\geq M)=\mu(\frac{|f|}{\phi\left(|f|\right)}\phi\left(|f|\right):|f|\geq M)\leq\varepsilon_M\mu(\phi\left(|f|\right):|f|\geq M)$$
$$\leq\varepsilon_M\mu(\phi\left(|f|\right))\leq K\varepsilon_M$$

and hence

$$\lim_{M\to\infty}\sup_{f\in\Lambda}\mu\left(|f|\,1_{|f|\geq M}\right)\leq\lim_{M\to\infty}K\varepsilon_M=0.$$

2. By assumption, $\varepsilon_M:=\sup_{f\in\Lambda}\mu\left(|f|\,1_{|f|\geq M}\right)\to 0$ as $M\to\infty$. Therefore we may choose $M_n\uparrow\infty$ such that

$$\sum_{n=0}^{\infty}(n+1)\,\varepsilon_{M_n}<\infty$$

where by convention $M_0:=0$. Now define ϕ so that $\phi(0)=0$ and

$$\phi'(x)=\sum_{n=0}^{\infty}(n+1)\,1_{(M_n,M_{n+1}]}(x),$$

i.e.

$$\phi(x)=\int_0^x\phi'(y)dy=\sum_{n=0}^{\infty}(n+1)\left(x\wedge M_{n+1}-x\wedge M_n\right).$$

By construction ϕ is continuous, $\phi(0)=0$, $\phi'(x)$ is increasing (so ϕ is convex) and $\phi'(x)\geq(n+1)$ for $x\geq M_n$. In particular

$$\frac{\phi(x)}{x}\geq\frac{\phi(M_n)+(n+1)x}{x}\geq n+1\text{ for }x\geq M_n$$

from which we conclude $\lim_{x\to\infty}\phi(x)/x=\infty$. We also have $\phi'(x)\leq(n+1)$ on $[0,M_{n+1}]$ and therefore

$$\phi(x)\leq(n+1)x\text{ for }x\leq M_{n+1}.$$

So for $f\in\Lambda$,

$$\begin{aligned}\mu\left(\phi(|f|)\right)&=\sum_{n=0}^{\infty}\mu\left(\phi(|f|)1_{(M_n,M_{n+1}]}(|f|)\right)\\&\leq\sum_{n=0}^{\infty}(n+1)\,\mu\left(|f|\,1_{(M_n,M_{n+1}]}(|f|)\right)\\&\leq\sum_{n=0}^{\infty}(n+1)\,\mu\left(|f|\,1_{|f|\geq M_n}\right)\leq\sum_{n=0}^{\infty}(n+1)\,\varepsilon_{M_n}\end{aligned}$$

and hence

$$\sup_{f\in\Lambda}\mu\left(\phi(|f|)\right)\leq\sum_{n=0}^{\infty}(n+1)\,\varepsilon_{M_n}<\infty.$$

■

Theorem 21.37 (Vitali Convergence Theorem). *(Folland 6.15) Suppose that* $1 \le p < \infty$. *A sequence* $\{f_n\} \subset L^p$ *is Cauchy iff*

1. $\{f_n\}$ *is* L^0 *– Cauchy,*
2. $\{|f_n|^p\}$ *– is uniformly integrable.*
3. *For all* $\varepsilon > 0$, *there exists a set* $E \in \mathcal{M}$ *such that* $\mu(E) < \infty$ *and* $\int_{E^c} |f_n|^p \, d\mu < \varepsilon$ *for all* n. *(This condition is vacuous when* $\mu(X) < \infty$.*)*

Proof. ($\Longrightarrow$) Suppose $\{f_n\} \subset L^p$ is Cauchy. Then (1) $\{f_n\}$ is L^0 – Cauchy by Lemma 21.14. (2) By completeness of L^p, there exists $f \in L^p$ such that $\|f_n - f\|_p \to 0$ as $n \to \infty$. By the mean value theorem,

$$||f|^p - |f_n|^p| \le p(\max(|f|, |f_n|))^{p-1} \, ||f| - |f_n|| \le p(|f| + |f_n|)^{p-1} \, ||f| - |f_n||$$

and therefore by Hölder's inequality,

$$\begin{aligned}\int ||f|^p - |f_n|^p| \, d\mu &\le p \int (|f| + |f_n|)^{p-1} \, ||f| - |f_n|| \, d\mu \le p \int (|f| + |f_n|)^{p-1} |f - f_n| d\mu \\ &\le p\|f - f_n\|_p \|(|f| + |f_n|)^{p-1}\|_q = p\| \, |f| + |f_n| \|_p^{p/q} \|f - f_n\|_p \\ &\le p(\|f\|_p + \|f_n\|_p)^{p/q} \|f - f_n\|_p\end{aligned}$$

where $q := p/(p-1)$. This shows that $\int ||f|^p - |f_n|^p| \, d\mu \to 0$ as $n \to \infty$.[3] By the remarks prior to Definition 21.32, $\{|f_n|^p\}$ is uniformly integrable. To verify (3), for $M > 0$ and $n \in \mathbb{N}$ let $E_M = \{|f| \ge M\}$ and $E_M(n) = \{|f_n| \ge M\}$. Then $\mu(E_M) \le \frac{1}{M^p} \|f\|_p^p < \infty$ and by the dominated convergence theorem,

$$\int_{E_M^c} |f|^p \, d\mu = \int |f|^p \, 1_{|f| < M} d\mu \to 0 \text{ as } M \to 0.$$

Moreover,

$$\left\|f_n 1_{E_M^c}\right\|_p \le \left\|f 1_{E_M^c}\right\|_p + \left\|(f_n - f) 1_{E_M^c}\right\|_p \le \left\|f 1_{E_M^c}\right\|_p + \|f_n - f\|_p . \tag{21.37}$$

So given $\varepsilon > 0$, choose N sufficiently large such that for all $n \ge N$, $\|f - f_n\|_p^p < \varepsilon$. Then choose M sufficiently small such that $\int_{E_M^c} |f|^p \, d\mu < \varepsilon$ and $\int_{E_M^c(n)} |f|^p \, d\mu < \varepsilon$ for all $n = 1, 2, \dots, N-1$. Letting $E := E_M \cup E_M(1) \cup \dots \cup E_M(N-1)$, we have

$$\mu(E) < \infty, \ \int_{E^c} |f_n|^p \, d\mu < \varepsilon \text{ for } n \le N - 1$$

and by Eq. (21.37)

[3] Here is an alternative proof. Let $h_n \equiv ||f_n|^p - |f|^p| \le |f_n|^p + |f|^p =: g_n \in L^1$ and $g \equiv 2|f|^p$. Then $g_n \xrightarrow{\mu} g$, $h_n \xrightarrow{\mu} 0$ and $\int g_n \to \int g$. Therefore by the dominated convergence theorem in Corollary 21.17, $\lim\limits_{n\to\infty} \int h_n \, d\mu = 0$.

$$\int_{E^c} |f_n|^p \; d\mu < (\varepsilon^{1/p} + \varepsilon^{1/p})^p \leq 2^p \varepsilon \text{ for } n \geq N.$$

Therefore we have found $E \in \mathcal{M}$ such that $\mu(E) < \infty$ and

$$\sup_n \int_{E^c} |f_n|^p \, d\mu \leq 2^p \varepsilon$$

which verifies (3) since $\varepsilon > 0$ was arbitrary. ($\Longleftarrow$) Now suppose$\{f_n\} \subset L^p$ satisfies conditions (1) - (3). Let $\varepsilon > 0$, E be as in (3) and

$$A_{mn} := \{x \in E | f_m(x) - f_n(x)| \geq \varepsilon\}.$$

Then

$$\|(f_n - f_m) \; 1_{E^c}\|_p \leq \|f_n 1_{E^c}\|_p + \|f_m \, 1_{E^c}\|_p < 2\varepsilon^{1/p}$$

and

$$\begin{aligned} \|f_n - f_m\|_p &= \|(f_n - f_m)1_{E^c}\|_p + \|(f_n - f_m)1_{E \setminus A_{mn}}\|_p \\ &\quad + \|(f_n - f_m)1_{A_{mn}}\|_p \\ &\leq \|(f_n - f_m)1_{E \setminus A_{mn}}\|_p + \|(f_n - f_m)1_{A_{mn}}\|_p + 2\varepsilon^{1/p}. \end{aligned} \tag{21.38}$$

Using properties (1) and (3) and $1_{E \cap \{|f_m - f_n| < \varepsilon\}} |f_m - f_n|^p \leq \; \varepsilon^p 1_E \in L^1$, the dominated convergence theorem in Corollary 21.17 implies

$$\|(f_n - f_m) \; 1_{E \setminus A_{mn}}\|_p^p = \int 1_{E \cap \{|f_m - f_n| < \varepsilon\}} \; |f_m - f_n|^p \underset{m,n \to \infty}{\longrightarrow} 0.$$

which combined with Eq. (21.38) implies

$$\limsup_{m,n \to \infty} \|f_n - f_m\|_p \leq \limsup_{m,n \to \infty} \|(f_n - f_m)1_{A_{mn}}\|_p + 2\varepsilon^{1/p}.$$

Finally

$$\|(f_n - f_m)1_{A_{mn}}\|_p \leq \|f_n 1_{A_{mn}}\|_p + \|f_m \; 1_{A_{mn}}\|_p \leq 2\delta(\varepsilon)$$

where

$$\delta(\varepsilon) := \sup_n \; \sup\{ \; \|f_n \; 1_E\|_p : E \in \mathcal{M} \; \ni \mu(E) \leq \varepsilon\}$$

By property (2), $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Therefore

$$\limsup_{m,n \to \infty} \|f_n - f_m\|_p \leq 2\varepsilon^{1/p} + 0 + 2\delta(\varepsilon) \to 0 \text{ as } \varepsilon \downarrow 0$$

and therefore $\{f_n\}$ is L^p-Cauchy. ■

Here is another version of Vitali's Convergence Theorem.

Theorem 21.38 (Vitali Convergence Theorem). *(This is problem 9 on p. 133 in Rudin.) Assume that $\mu(X) < \infty$, $\{f_n\}$ is uniformly integrable, $f_n \to f$ a.e. and $|f| < \infty$ a.e., then $f \in L^1(\mu)$ and $f_n \to f$ in $L^1(\mu)$.*

Proof. Let $\varepsilon > 0$ be given and choose $\delta > 0$ as in the Eq. (21.32). Now use Egoroff's Theorem 21.18 to choose a set E^c where $\{f_n\}$ converges uniformly on E^c and $\mu(E) < \delta$. By uniform convergence on E^c, there is an integer $N < \infty$ such that $|f_n - f_m| \le 1$ on E^c for all $m, n \ge N$. Letting $m \to \infty$, we learn that

$$|f_N - f| \le 1 \text{ on } E^c.$$

Therefore $|f| \le |f_N| + 1$ on E^c and hence

$$\begin{aligned}\mu(|f|) &= \mu(|f| : E^c) + \mu(|f| : E) \\ &\le \mu(|f_N|) + \mu(X) + \mu(|f| : E).\end{aligned}$$

Now by Fatou's lemma,

$$\mu(|f| : E) \le \liminf_{n\to\infty} \mu(|f_n| : E) \le 2\varepsilon < \infty$$

by Eq. (21.32). This shows that $f \in L^1$. Finally

$$\begin{aligned}\mu(|f - f_n|) &= \mu(|f - f_n| : E^c) + \mu(|f - f_n| : E) \\ &\le \mu(|f - f_n| : E^c) + \mu(|f| + |f_n| : E) \\ &\le \mu(|f - f_n| : E^c) + 4\varepsilon\end{aligned}$$

and so by the Dominated convergence theorem we learn that

$$\limsup_{n\to\infty} \mu(|f - f_n|) \le 4\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary this completes the proof. ■

Theorem 21.39 (Vitali again). *Suppose that $f_n \to f$ in μ measure and Eq. (21.34) holds, then $f_n \to f$ in L^1.*

Proof. This could of course be proved using 21.38 after passing to subsequences to get $\{f_n\}$ to converge a.s. However I wish to give another proof. First off, by Fatou's lemma, $f \in L^1(\mu)$. Now let

$$\phi_K(x) = x1_{|x|\le K} + K1_{|x|>K}.$$

then $\phi_K(f_n) \xrightarrow{\mu} \phi_K(f)$ because $|\phi_K(f) - \phi_K(f_n)| \le |f - f_n|$ and since

$$|f - f_n| \le |f - \phi_K(f)| + |\phi_K(f) - \phi_K(f_n)| + |\phi_K(f_n) - f_n|$$

we have that

$$\begin{aligned}\mu|f - f_n| &\le \mu\,|f - \phi_K(f)| + \mu|\phi_K(f) - \phi_K(f_n)| + \mu\,|\phi_K(f_n) - f_n| \\ &= \mu(|f| : |f| \ge K) + \mu|\phi_K(f) - \phi_K(f_n)| + \mu(|f_n| : |f_n| \ge K).\end{aligned}$$

Therefore by the dominated convergence theorem

$$\limsup_{n\to\infty} \mu|f - f_n| \le \mu(|f| : |f| \ge K) + \limsup_{n\to\infty} \mu(|f_n| : |f_n| \ge K).$$

This last expression goes to zero as $K \to \infty$ by uniform integrability. ■

21.6 Exercises

Definition 21.40. *The **essential range** of* f, $\mathrm{essran}(f)$, *consists of those* $\lambda \in \mathbb{C}$ *such that* $\mu(|f-\lambda| < \varepsilon) > 0$ *for all* $\varepsilon > 0$.

Definition 21.41. *Let* (X, τ) *be a topological space and* ν *be a measure on* $\mathcal{B}_X = \sigma(\tau)$. *The **support** of* ν, $\mathrm{supp}(\nu)$, *consists of those* $x \in X$ *such that* $\nu(V) > 0$ *for all open neighborhoods,* V, *of* x.

Exercise 21.3. Let (X, τ) be a second countable topological space and ν be a measure on $\mathcal{B}_X$ – the Borel σ – algebra on X. Show

1. $\mathrm{supp}(\nu)$ is a closed set. (This is actually true on all topological spaces.)
2. $\nu(X \setminus \mathrm{supp}(\nu)) = 0$ and use this to conclude that $W := X \setminus \mathrm{supp}(\nu)$ is the largest open set in X such that $\nu(W) = 0$. **Hint:** $\mathcal{U} \subset \tau$ be a countable base for the topology τ. Show that W may be written as a union of elements from $V \in \mathcal{V}$ with the property that $\mu(V) = 0$.

Exercise 21.4. Prove the following facts about $\mathrm{essran}(f)$.

1. Let $\nu = f_*\mu := \mu \circ f^{-1}$ – a Borel measure on $\mathbb{C}$. Show $\mathrm{essran}(f) = \mathrm{supp}(\nu)$.
2. $\mathrm{essran}(f)$ is a closed set and $f(x) \in \mathrm{essran}(f)$ for almost every x, i.e. $\mu(f \notin \mathrm{essran}(f)) = 0$.
3. If $F \subset \mathbb{C}$ is a closed set such that $f(x) \in F$ for almost every x then $\mathrm{essran}(f) \subset F$. So $\mathrm{essran}(f)$ is the smallest closed set F such that $f(x) \in F$ for almost every x.
4. $\|f\|_\infty = \sup\{|\lambda| : \lambda \in \mathrm{essran}(f)\}$.

Exercise 21.5. Let $f \in L^p \cap L^\infty$ for some $p < \infty$. Show $\|f\|_\infty = \lim_{q\to\infty} \|f\|_q$. If we further assume $\mu(X) < \infty$, show $\|f\|_\infty = \lim_{q\to\infty} \|f\|_q$ for all measurable functions $f : X \to \mathbb{C}$. In particular, $f \in L^\infty$ iff $\lim_{q\to\infty} \|f\|_q < \infty$. **Hints:** Use Corollary 21.23 to show $\limsup_{q\to\infty} \|f\|_q \le \|f\|_\infty$ and to show $\liminf_{q\to\infty} \|f\|_q \ge \|f\|_\infty$, let $M < \|f\|_\infty$ and make use of Chebyshev's inequality.

Exercise 21.6. Prove Eq. (21.21) in Corollary 21.23. (Part of Folland 6.3 on p. 186.) **Hint:** Use the inequality, with $a, b \ge 1$ with $a^{-1} + b^{-1} = 1$ chosen appropriately,

$$st \le \frac{s^a}{a} + \frac{t^b}{b},$$

(see Lemma 5.5 for Eq. (21.16)) applied to the right side of Eq. (21.20).

Exercise 21.7. Complete the proof of Proposition 21.22 by showing $(L^p + L^r, \|\cdot\|)$ is a Banach space. (Part of Folland 6.4 on p. 186.)

Exercise 21.8. Folland 6.5 on p. 186.

Exercise 21.9. By making the change of variables, $u = \ln x$, prove the following facts:

$$\int_0^{1/2} x^{-a} |\ln x|^b \, dx < \infty \iff a < 1 \text{ or } a = 1 \text{ and } b < -1$$
$$\int_2^{\infty} x^{-a} |\ln x|^b \, dx < \infty \iff a > 1 \text{ or } a = 1 \text{ and } b < -1$$
$$\int_0^{1} x^{-a} |\ln x|^b \, dx < \infty \iff a < 1 \text{ and } b > -1$$
$$\int_1^{\infty} x^{-a} |\ln x|^b \, dx < \infty \iff a > 1 \text{ and } b > -1.$$

Suppose $0 < p_0 < p_1 \leq \infty$ and m is Lebesgue measure on $(0, \infty)$. Use the above results to manufacture a function f on $(0, \infty)$ such that $f \in L^p((0, \infty), m)$ iff (a) $p \in (p_0, p_1)$, (b) $p \in [p_0, p_1]$ and (c) $p = p_0$.

Exercise 21.10. Folland 6.9 on p. 186.

Exercise 21.11. Folland 6.10 on p. 186. Use the strong form of Theorem 19.38.

Exercise 21.12. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ – finite measure spaces, $f \in L^2(\nu)$ and $k \in L^2(\mu \otimes \nu)$. Show

$$\int |k(x, y) f(y)| \, d\nu(y) < \infty \text{ for } \mu - \text{a.e. } x.$$

Let $Kf(x) := \int_Y k(x, y) f(y) d\nu(y)$ when the integral is defined. Show $Kf \in L^2(\mu)$ and $K : L^2(\nu) \to L^2(\mu)$ is a bounded operator with $\|K\|_{op} \leq \|k\|_{L^2(\mu \otimes \nu)}$.

Exercise 21.13. Folland 6.27 on p. 196. **Hint**: Theorem 21.28.

Exercise 21.14. Folland 2.32 on p. 63.

Exercise 21.15. Folland 2.38 on p. 63.