Measure Theory and Integration

Lecture Notes
Changyu Xia

1. Introduction

In order to know why we need a theory of measures and a new way to define integrals (Lebesgue integrals), we look at the problems arised in Riemann integration which uses rectangles to approximate areas under graphs. Let us recall the definition of it. For a bounded function f defined on $[a,b]$, and for any partition $P : a = x_0 < x_1 < \cdots x_n = b$, the lower Riemann sum L_P and the upper Riemann sum U_P are given by

$$L_P = \sum_{j=1}^{n} (x_j - x_{j-1}) \inf_{[x_{j-1},x_j]} f$$

and

$$U_P = \sum_{j=1}^{n} (x_j - x_{j-1}) \sup_{[x_{j-1},x_j]} f.$$

We define the Riemann lower integral

$$\underline{\int_a^b} f(x)dx = \sup_P L_P$$

and the Riemann upper integral

$$\overline{\int_a^b} f(x)dx = \inf_P U_P$$

We call f Riemann integrable over $[a,b]$ if $\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx$ and denote this commom value by $\int_a^b f(x)dx$. An important result regarding Riemann integration says that every continuous real-valued function on each closed bounded interval is Riemann integrable. Another useful theorem in Riemann integration says that if f_n is a sequence of Riemann integrable functions on $[a,b]$ converging uniformly on $[a,b]$ to a function $f : [a,b] \to \mathbb{R}$, then f is Riemann integrable and $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$.

The Riemann integral works well for students studying calculus around the world but has several deficiencies. It does not handle unbounded functions and functions with many discontinuities and does not work well with limits. Also some quite simple functions are not Riemann integrable. A typical example is the Dirichlet function defined by

$$D(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{Q}^c, \end{cases} \tag{1.1}$$

which is not Riemann integrable over any $[a, b]$ since $\underline{\int_a^b} D(x)dx = 0$ and $\overline{\int_a^b} D(x)dx = b - a$.

For nonnegative Riemann integrable functions $\{f_n\}$, $\int_a^b f_n dx \geq 0$, so $\sum_{n=1}^{\infty} \int_a^b f_n dx$ always exists (maybe ∞). It is natural to expect

$$\sum_{n=1}^{\infty} \int_a^b f_n dx = \int_a^b \sum_{n=1}^{\infty} f_n dx.$$

But it is not true in general, as one can see that $D(x) = \sum_{n=1}^{\infty} f_n(x)$, where, for an enumeration $\mathbb{Q} = \{r_1, r_2, r_3, ...,\}$,

$$f_n(x) = \begin{cases} 1, & x = r_n, \\ 0, & \text{otherwise}, \end{cases}$$

Each f_n is Riemann integrable over any [a,b] and $\int_a^b f_n dx = 0$.

Also, for an increasing sequence of nonnegative Riemann integrable functions $0 \leq g_1 \leq g_2 \leq$ such that $\lim_{n\to\infty} g_n(x) = g(x)$, it is natural to expect that $\lim_{n\to\infty} \int_a^b f_n dx = \int_a^b f dx$. However, this is not true for Riemann integrals, as one can see that $D(x) = \lim_{n\to\infty} g_n(x)$, where

$$g_n(x) = \begin{cases} 1, & x = r_1, r_2, \cdots, r_n, \\ 0, & \text{otherwise}. \end{cases}$$

Each g_n is Riemann integrable over any $[a, b]$ and $\int_a^b g_n dx = 0$.

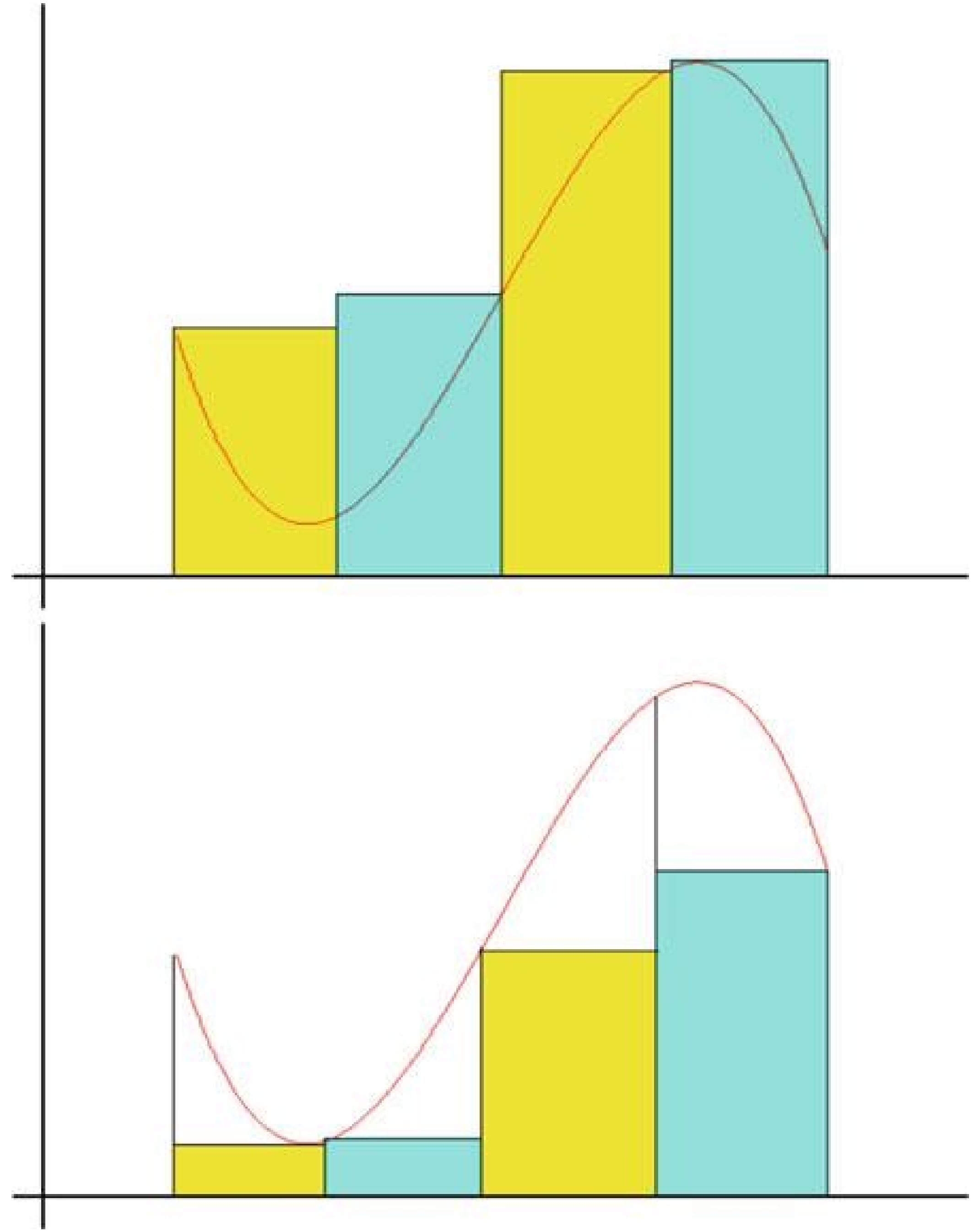
Because analysis relies heavily upon limits, a good theory of integration should allow for interchange of limits and integrals, at least when the functions are appropriately bounded. Thus the previous example points out a serious deficiency in Riemann integration and a new theory of integration is needed.

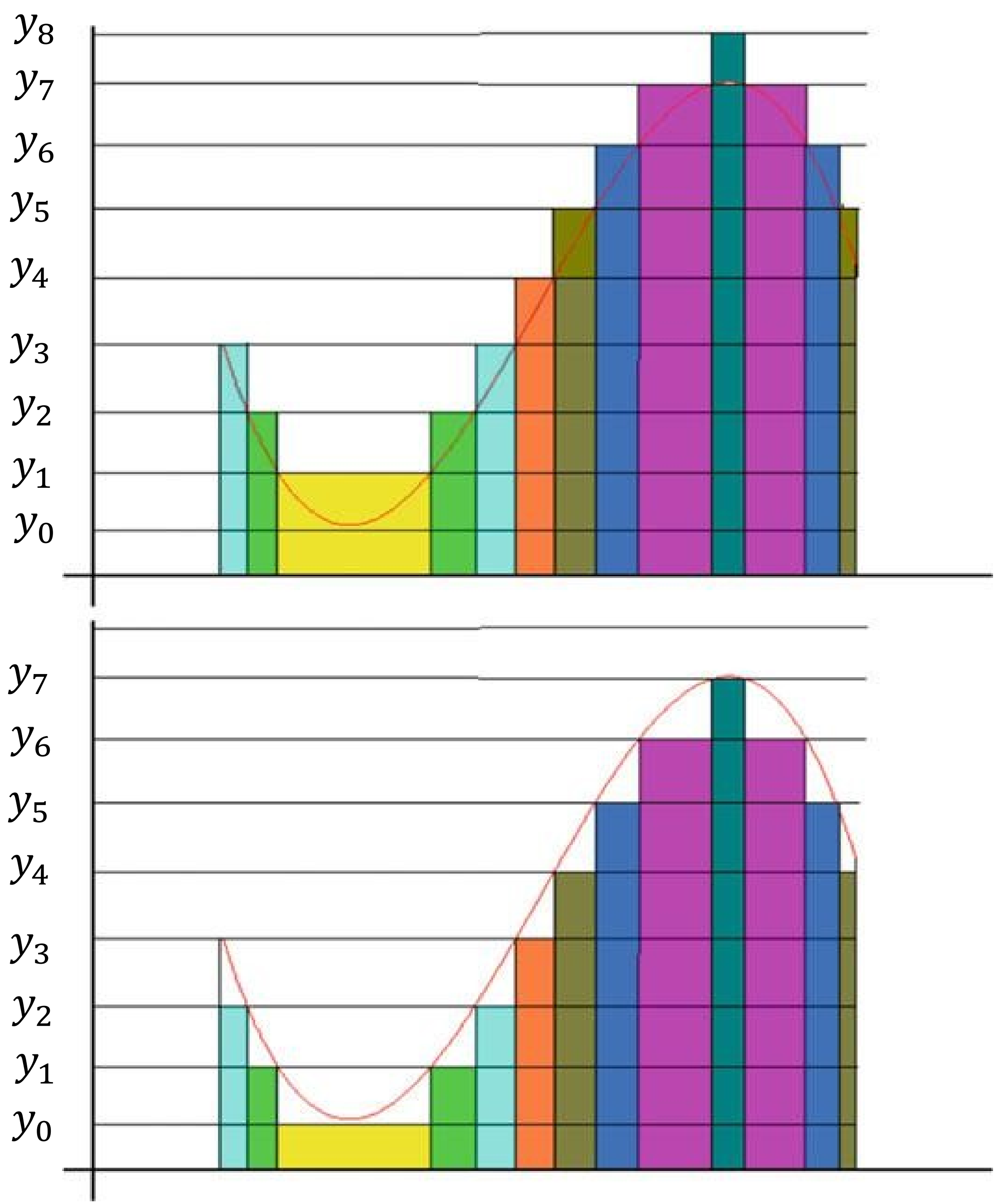
In the early twentieth century, Lebesgue defined a new integral, which he used to give very useful answers to questions of the sort discussed above. Instead of taking a partition of the x-axis, one can take a partition of the y-axis. For each subinterval $(y_{i-1}, y_i]$, we consider the set $E_i = \{x \in [a, b] : y_{i-1} < f(x) \leq y_i\}$. Since on E_i, $y_{i-1} < f(x) \leq y_i$, the "area" of the region under f is bounded below by $y_{i-1} l(E_i)$ and above by $y_i l(E_i)$, where $l(E_i)$ is the *size*, in a sense still to be defined, of the set E_i. Assuming that the values of f are between y_0 and y_n, similar to the Riemann integrals, one can form lower "Lebesgue" sum

$$\sum_{i=1}^{n} y_{i-1} l(E_i)$$

and upper "Lebesgue" sum

$$\sum_{i=1}^{n} y_i l(E_i).$$




y_8
y_7
y_6
y_5
y_4
y_3
y_2
y_1
y_0
y_7
y_6
y_5
y_4
y_3
y_2
y_1
y_0

Since the difference between the upper and lower "Lebesgue" sums is at most $(b-a)\max_i(y_i - y_{i-1})$, and goes to zero as $\max_i(y_i - y_{i-1}) \to 0$, such a definition of integral works for any bounded function, provided that one can properly define "length" of E_i. Thus comes the terminology of a "measure" of a set. In 1902, Lebesgue for the first time generalized the concepts "length" and "area" of $\mathbb{R}^2$ to the Lebesgue measure on Borel sets and also defined the Lebesgue integral of Lebesgue measurable functions on Lebesgue measurable sets. The idea of Lebesgue brought great developments to analysis. In 1915, Frécht defined measures and signed measures on abstract measurable spaces. Measure theory now becomes an important area of modern Mathematics and is closed related to many other subareas. We will start with the measurable spaces.

2. Algebras, σ-algebras, measurable spaces

We need to know how to measure the "size" or "volume" of subsets of a space X before we can integrate functions on X. We're familiar with volume of subsets in $\mathbb{R}^n$. What about more general spaces X ? We need a *measure* function $\mu : \{subsets\ of\ X\} \to [0,\infty]$. For technical reasons, a measure will not be defined on all subsets of X, but instead a certain collection of subsets of X called a σ-algebra.

Denote by $\mathcal{P}(X)$ the set of all subsets of X. Let $\{A_n\}$ be a sequence in $\mathcal{P}(X)$. Then the following De Morgan identity holds,

$$(\cup_{k=1}^{\infty} A_k)^c = \cap_{k=1}^{\infty} A_k^c.$$

We define

$$\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{m=n}^{\infty} A_m, \quad \liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m.$$

It can be checked that $\limsup_{n\to\infty} A_n$ (respectively $\liminf_{n\to\infty} A_n$) consists of those elements of X that belong to infinitely many A_n (respectively that belong to all but finitely many A_n).

Definition 2.1 *A non-empty subset $\mathcal{A} \subset \mathcal{P}(X)$ is said to be an algebra if*
i) $\emptyset \in \mathcal{A}$,
ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.
iii) $A_1, ..., A_n \in \mathcal{A} \Rightarrow \cup_{i=1}^n A_i \in \mathcal{A}$.

From $\cap_{i=1}^n A_i = (\cup_{i=1}^n A_i^c)^c$, we know that an algebra is also closed under finite intersections.

Contrast with a *topology* $\tau \subset \mathcal{P}(X)$, which satisfies
i) $\emptyset \in \tau$ and $X \in \tau$;
ii) if $U_i \in \tau (i = 1, ..., n)$, then $\cap_{i=1}^n U_i \in \tau$;
iii) if $U_\alpha (\alpha \in \mathcal{I})$ is an arbitrary collection in τ, then $\cup_{\alpha\in\mathcal{I}} U_\alpha \in \tau$.

Let $\mathcal{K} \subset \mathcal{P}(X)$. As the intersection of any family of algebras is still an algebra, the minimal algebra including $\mathcal{K}$ (that is the intersection of all algebras including $\mathcal{K}$) is well defined, and called the algebra generated by $\mathcal{K}$.

Definition 2.2 *A non empty subset $\mathcal{A} \subset \mathcal{P}(X)$ is said to be a $\sigma-algebra$ if* *i) $\mathcal{A}$ is an algebra;*

(ii) if $\{A_n\}$ is a sequence of elements of $\mathcal{A}$ then $\cup_{n=1}^{\infty} A_i \in \mathcal{A}$.

In this case, we call $(X, \mathcal{A})$ a measurable space and elements of $\mathcal{A}$ are called $\mathcal{A}$-measurable sets or simply measurable sets.

It follows from $\cap_{i=1}^{\infty} A_i = (\cup_{i=1}^{\infty} A_i^c)^c$ that a σ-algebra is also closed under countable intersections. That is, if $\mathcal{A}$ is a σ-algebra and $\{A_n\} \subset \mathcal{A}$, then $\cap_{n=1}^{\infty} A_n \in \mathcal{A}$.

Let $\mathcal{K} \subset \mathcal{P}(X)$. As the intersection of any family of σ-algebras is still a σ-algebra, the minimal σ-algebra including $\mathcal{K}$ (that is the intersection of all σ-algebras including $\mathcal{K}$) is well defined, and called the σ-algebra generated by $\mathcal{K}$. It is denoted by $\sigma(\mathcal{K})$.

Examples:

a) $\{\emptyset, X\}$ and $\mathcal{P}(X)$ are σ-algebras.

b) Let X be an infinite set, and let $\mathcal{B}$ be the collection of all finite subsets of X. Then $\mathcal{B}$ does not contain X and so is not an algebra.

c) Let X be an infinite set, and let $\mathcal{A}$ be the collection of all subsets A of X such that either A or A^c is finite. Then $\mathcal{A}$ is an algebra on X but is not closed under the formation of countable unions; hence it is not a σ-algebra.

d) Let X be an uncountable set, and let $\mathcal{A}$ be the collection of all countable (i.e., finite or countably infinite) subsets of X. Then $\mathcal{A}$ does not contain X; hence it is not an algebra.

e) Let X be a set, and let $\mathcal{A}$ be the collection of all subsets A of X such that either A or A^c is countable. Then $\mathcal{A}$ is a σ-algebra. The non-trivial part we need to show is that if $\{A_i\} \subset \mathcal{A}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$. Suppose $A_1, A_2, ...$ are each in $\mathcal{A}$. If all of the $A_i's$ are countable, then $\cup_{i=1}^{\infty} A_i$ is countable, and so is in $\mathcal{A}$. If $A_{i_0}^c$ is countable for some i_0, then

$$(\cup_{i=1}^{\infty} A_i)^c = \cap_{i=1}^{\infty} A_i^c \subset A_{i_0}^c$$

is countable, and again $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$. Hence, $\mathcal{A}$ is a σ-algebra.

f) Let $X = [0, 1]$ and let $\mathcal{A} = \{\emptyset, X, [0, \frac{1}{2}), [\frac{1}{2}, 1]\}$. Then $\mathcal{A}$ is a σ-algebra.

g) Let $X = \{1, 2, 3\}$ and let $\mathcal{A} = \{\emptyset, X, \{1\}, \{2, 3\}\}$. Then $\mathcal{A}$ is a σ-algebra.

Definition 2.3 (Borel σ-algebra). *If X is a topological space, the σ-algebra generated by all open subsets of X is called the Borel σ-algebra of X and it is denoted by $\mathcal{B}(X)$.*

In the case when $X = \mathbb{R}$ the Borel σ-algebra has a particularly simple class of generators.

Proposition 2.4 *The σ-algebra $\mathcal{B}(\mathbb{R})$ of Borel subsets of $\mathbb{R}$ is generated by each of the following collections of sets:*

(a) the collection of all closed subsets of $\mathbb{R}$;

(b) the collection of all subintervals of $\mathbb{R}$ of the form $(-\infty, b]$;

(c) the collection of all subintervals of $\mathbb{R}$ of the form $(a, b]$.

Proof. Let $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$ be the σ-algebras generated by the collections of sets in parts (a), (b), and (c) of the proposition. We will show that $\mathcal{B}(\mathbb{R}) \supset \mathcal{B}_1 \supset \mathcal{B}_2 \supset \mathcal{B}_3$ and then $\mathcal{B}_3 \supset \mathcal{B}(\mathbb{R})$; this will establish the proposition. Since $\mathcal{B}(\mathbb{R})$ includes the family of open subsets of $\mathbb{R}$ and is closed under complementation, it includes the family of closed subsets of $\mathbb{R}$; thus it includes the σ-algebra generated by the closed subsets of $\mathbb{R}$, namely $\mathcal{B}_1$. The sets of the form $(-\infty, b]$ are closed and so belong to $\mathcal{B}_1$; consequently $\mathcal{B}_1 \supset \mathcal{B}_2$. Since $(a,b] = (-\infty,b] \cap (-\infty,a]^c$, each set of the form $(a,b]$ belongs to $\mathcal{B}_2$; thus $\mathcal{B}_2 \supset \mathcal{B}_3$. Finally, note that each open interval of $\mathbb{R}$ is the union of a sequence of sets of the form $(a,b]$ and that each open subset of $\mathbb{R}$ is the union of a sequence of open intervals. Thus each open subset of $\mathbb{R}$ belongs to $\mathcal{B}_3$, and so $\mathcal{B}_3 \supset \mathcal{B}(\mathbb{R})$.

Proposition 2.5 *The Borel σ-algebra on $\mathbb{R}^d$ is generated by each of the following collections of sets:*
(a) the collection of all closed subsets of $\mathbb{R}^d$;
(b) the collection of all closed half-spaces in $\mathbb{R}^d$ that have the form $\{(x_1,...,x_d) : x_i \le b\}$ for some index i and some b in $\mathbb{R}$;
(c) the collection of all rectangles in $\mathbb{R}^d$ that have the form $\{(x_1,...,x_d) : a_i < x_i \le b_i \ for\ i = 1,...,d\}$.

Proof. This proposition can be proved with essentially the argument that was used for Proposition 2.4, and so most of the proof is omitted. To see that the σ-algebra generated by the rectangles of part (c) is included in the σ-algebra generated by the half-spaces of part (b), note that each strip that has the form $\{(x_1,...,x_d) : a < x_i \le b\}$ for some i is the difference of two of the half-spaces in part (b) and that each of the rectangles in part (c) is the intersection of d such strips. □

- F_σ (“F-sigma”): countable union of closed sets.
- G_δ (“G-delta”): countable intersection of open sets.

Exercises 1

1. Let $E \subset X$ be any set and let $\mathcal{A}$ be some σ-algebra on X. Show that

$$\mathcal{A}_E = \{E \cap A : A \subset \mathcal{A}\}$$

is a σ-algebra.

2. Let $\mathcal{A}$ be an algebra of subsets of X and $A,\ B \notin \mathcal{A}$. Is the same true of $A \cup B$?

3. Let X be an infinite set. Describe the σ-algebra generated by the singletons of X.

4. Show that the union of algebras of subsets of X is not necessarily an algebra. How about the union of an increasing collection of algebras ?

5. Let $\mathcal{A}_1, \cdots, \mathcal{A}_n$ be σ-algebras of subsets of X such that $\mathcal{A} = \cup_{k=1}^n \mathcal{A}_k$ is an algebra. Show that $\mathcal{A}$ is a σ-algebra.

6. Let $\mathcal{A}$ be a σ-algebra of subsets of X that contains countably infinite nonempty pairwise disjoints sets. Prove that $\mathcal{A}$ is uncountable.

7. Does there exist an infinite σ-algebra $\mathcal{A}$ with countably many sets ?

8. An algebra $\mathcal{A}$ is a σ-algebra iff $\mathcal{A}$ is closed under countable increasing unions (i.e., if $\{E_j\}_{j=1}^{\infty} \subset \mathcal{A}$ and $E_1 \subset E_2 \subset \cdots$, then $\cup_{j=1}^{\infty} E_j \in \mathcal{A}$).

9. Let $f : X \to Y$ be a map and $\mathcal{A}$ be a σ-algebra on Y. Then

$$\mathcal{A}' = \{f^{-1}(B) : B \in \mathcal{A}\}$$

is a σ-algebra on X.

10. If $\mathcal{A}$ is the σ-algebra generated by $\mathcal{M}$, then $\mathcal{A}$ is the union of the σ-algebras generated by $\mathcal{F}$ as $\mathcal{F}$ ranges over all countable subsets of $\mathcal{M}$. (Hint: Show that the latter object is a σ-algebra.)

11. Let $\mathcal{A}$ be the collection of all subsets of $\mathbb{R}$ that are unions of finitely many intervals of the form $(a, b], (a, \infty), or(-\infty, b]$. Show that $\mathcal{A}$ is an algebra but is not a σ-algebra.

12. Let $A \subset \mathbb{R}^n$ be a Borel set. Show that for $h \in \mathbb{R}^n$ and $t \in \mathbb{R}$ the sets $A + h := \{a + h : a \in A\},\ tA := \{ta : a \in A\}$ are Borel as well.

13. Let X be a set and let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra with finite cardinality. Show that its cardinality is equal to 2^n for some integer $n \geq 1$.

3. Measures

In this section we give the definition of a measure, some examples, and the basic properties of a measure.

Definition 3.1 Let $(X, \mathcal{A})$ be a measurable space. A measure on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

(1) $\mu(\emptyset) = 0$;

(2) if $A_i \in \mathcal{A}, i = 1, ...$ are pairwise disjoint, then

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

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

Condition (2) in Definition 3.1 is known as *countable additivity* or *σ-additivity*. We say that a set function is *finitely additive* if $\mu(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$ whenever $A_1, ..., A_n$ are in $\mathcal{A}$ and are pairwise disjoint.

Examples

a) Let X be any set, $\mathcal{A}$, the collection of all subsets of X and let $\mu(A)$= the number of elements in A if A is finite and $\mu(A) = \infty$ if A is infinite. μ is called *counting measure*.

b) Let $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise. This measure is called *Dirac-measure at x*.

c) Let X be an uncountable set and $\mathcal{C}$ the collection of those subsets that are either countable or the complement of a countable set. Define $\mu : \mathcal{C} \to \{0, 1\}$ by $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A is uncountable. Then $(X, \mathcal{C}, \mu)$ is a measure space.

d) Let $\mathcal{L}$ be the collection of Lebesgue measurable sets of real numbers and m be the Lebesgue measure; the $(\mathbb{R}, \mathcal{L}, m)$ is a measure space. Also, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is a measure space.

e) Let X be an infinite set and define $\mu(A) = 0$ if A is countable and $\mu(A) = \infty$ otherwise. Then μ is a measure on $\mathcal{P}(X)$.

f) Let X be a set that has at least two members, and let $\mathcal{A} = \mathcal{P}(X)$. Define a function $\mu : \mathcal{A} \to [0, \infty]$ by letting $\mu(A)$ be 1 if $A \neq \emptyset$ and letting $\mu(A)$ be 0 if $A = \emptyset$. Then μ is not a measure.

Let $\{A_i\}_{i=1}^{\infty} \subset \mathcal{P}(X)$. If $A_1 \subset A_2 \subset \cdots$, $A = \cup_{i=1}^{\infty} A_i$, we write $A_i \uparrow A$. Similarly, if $A_1 \supset A_2 \supset \cdots$, $A = \cap_{i=1}^{\infty} A_i$, we write $A_i \downarrow A$.

Proposition 3.3 *Let $(X, \mathcal{A}, \mu)$ be a measure space. The following hold:*
*i) (**Monotonicity**) If $A, B \in \mathcal{A}$ with $A \subset B$, then $\mu(A) \leq \mu(B)$;*
*ii) (**Subadditivity**) If $A_i \in \mathcal{A}$ and $A = \cup_{i=1}^{\infty} A_i$, then $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$;*
*iii) (**Continuity from below**) If $A_i \in \mathcal{A}$ and $A_i \uparrow A$, then $\mu(A) = \lim_{i\to\infty} \mu(A_i)$;*
*iv) (**Continuity from above**) Let $A_i \in \mathcal{A}$ and $A_i \downarrow A$. If $\mu(A_1) < \infty$, then $\mu(A) = \lim_{i\to\infty} \mu(A_i)$.*

Proof. i) If $A \subset B$, then $\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$.

ii) Setting $B_1 = A_1$ and $B_j = A_j - (\cup_{i=1}^{j-1} A_i)$ for $j > 1$. Then the B_j's are disjoint and $\cup_{j=1}^{\infty} B_j = \cup_{j=1}^{\infty} A_j$. Therefore, by the σ-additivity of μ and (i),

$$\mu(\cup_{i=1}^{\infty} A_i) = \mu(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

iii) Setting $A_0 = \emptyset$, we have

$$\mu(\cup_{i=1}^{\infty} A_i) = \mu(\cup_{i=1}^{\infty}(A_i - A_{i-1})) = \sum_{i=1}^{\infty} \mu(A_i - A_{i-1}) = \lim_{i\to\infty} \mu(A_i).$$

iv) Let $B_j = A_1 - A_j$; then $\mu(A_1) = \mu(A_j) + \mu(B_j)$ and $B_j \uparrow A_1 - (\cap_{i=1}^{\infty} A_i)$. By (iii),

$$\mu(A_1) - \mu(\cap_{i=1}^{\infty} A_i) = \mu(A_1 - (\cap_{i=1}^{\infty} A_i)) = \lim_{i\to\infty} \mu(B_i) = \mu(A_1) - \lim_{i\to\infty} \mu(A_i).$$

$\square$

Remark 3.3 To see that $\mu(A_1) < \infty$ is necessary in Proposition 3.2, let $X = \mathbb{N}$, μ be the counting measure and $A_i = \{i, i+1, ...\}$. Then $A_i \downarrow \emptyset$ but $\mu(A_i) = \infty$ for all i.

Definition 3.4 A measure μ is a finite measure if $\mu(X) < \infty$. A measure μ is σ-finite if there exist sets $E_i \in \mathcal{A}$ for $i = 1, 2, ...$ such that $\mu(E_i) < \infty$ for each i and $X = \cup_{i=1}^{\infty} E_i$. If μ is a finite measure, then $(X, \mathcal{A}, \mu)$ is called a finite measure space, and similarly, if μ is a σ-finite measure, then $(X, \mathcal{A}, \mu)$ is called a σ-finite measure space.

Suppose $(X, \mathcal{A}, \mu)$ is σ-finite so that $X = \cup_{i=1}^{\infty} E_i$ with $\mu(E_i) < \infty$ and $E_i \in \mathcal{A}$ for each i. If we let $F_n = \cup_{i=1}^{n} E_i$, then $\mu(F_n) < \infty$ for each n and $F_n \uparrow X$. Therefore there is no loss of generality in supposing the sets E_i in Definition 3.4 are increasing.

For a measure space $(X, \mathcal{A}, \mu)$ and a measurable subset E of X, we say that a property holds *almost everywhere on* E, or it holds *for almost all* x *in* E, provided it holds on $E - E_0$, where E_0 is a measurable subset of E for which $\mu(E_0) = 0$.

The Borel-Cantelli Lemma *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and* $\{E_k\}_{k=1}^{\infty}$ *a countable collection of measurable sets for which* $\sum_{i=1}^{\infty} \mu(E_k) < \infty$. *Then almost all* x *in* X *belong to a finite number of the* E_k*'s.*

Proof For each n, by the countable monotonicity of μ, $\mu(\cup_{k=n}^{\infty} E_k) \le \sum_{k=n}^{\infty} \mu(E_k)$. Hence, by the continuity of μ,

$$\mu(\cap_{n=1}^{\infty}(\cup_{k=n}^{\infty} E_k)) = \lim_{n\to\infty} \mu(\cup_{k=n}^{\infty} E_k) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(E_n) = 0.$$

Observe that $\cap_{n=1}^{\infty}(\cup_{k=n}^{\infty} E_k)$ is the set of all points in X that belong to an infinite number of the E_k's. $\square$

Definition 3.5 A measure space $(X, \mathcal{A}, \mu)$ is said to be *complete* if

$$N \in \mathcal{A},\ \mu(N) = 0,\ E \subset N \Rightarrow E \in \mathcal{A}.$$

Example. Let $X = \{a, b, c\}$. Then $\mathcal{A} = \{\emptyset, \{a\}, \{b, c\}, X\}$ is a σ-algebra of subsets of X. If we define a set function μ on $\mathcal{A}$ by setting $\mu(\emptyset) = 0, \mu(\{a\}) = 1, \mu(\{b, c\}) = 0$, and $\mu(X) = 1$, then μ is a measure on $\mathcal{A}$. The set $\{b, c\}$ is a null set in the measure space $(X, \mathcal{A}, \mu)$, but its subset $\{b\}$ is not a member of $\mathcal{A}$. Therefore $(X, \mathcal{A}, \mu)$ is not a complete measure space.

Theorem 3.6 *Let* $(X, \mathcal{A}, \mu)$ *be a measure space and define*

$$\mathcal{A}^* = \{E \subset X : \exists\ A, B \in \mathcal{A} \text{ such that } A \subset E \subset B \text{ and } \mu(B - A) = 0\}.$$

Then the following holds.

(i) $\mathcal{A}^*$ *is a* σ*-algebra and* $\mathcal{A} \subset \mathcal{A}^*$.

(ii) There exists a unique measure μ^* *on* $\mathcal{A}^*$ *such that*

$$\mu^*|_{\mathcal{A}} = \mu.$$

(iii) The triple $(X, \mathcal{A}^*, \mu^*)$ *is a complete measure space. It is called the completion of* $(X, \mathcal{A}, \mu)$.

Proof. First, $X \in \mathcal{A}^*$ and $\mathcal{A} \subset \mathcal{A}^*$. Second, let $E \in \mathcal{A}^*$ and choose $A, B \in \mathcal{A}$ such that $A \subset E \subset B$. Then $B^c \subset E^c \subset A^c$ and $A^c - B^c = A^c \cap B = B - A$. Hence $\mu(A^c - B^c) = 0$ and so $E^c \in \mathcal{A}^*$. Third, let $E_i \in \mathcal{A}^*,\ i \in \mathbb{N}$ and choose $A_i, B_i \in \mathcal{A}$ such that $A_i \subset E_i \subset B_i$ and $\mu(B_i - A_i) = 0$. Define

$$A = \cup_{i=1}^{\infty} A_i,\ E = \cup_{i=1}^{\infty} E_i,\ B = \cup_{i=1}^{\infty} B_i.$$

Then $A \subset E \subset B$ and $B - A = \cup_i (B_i - A) \subset \cup_i (B_i - A_i)$. Hence

$$\mu(B - A) \le \sum_{i=1}^{\infty} \mu(B_i - A_i) = 0$$

and this implies $E \in \mathcal{A}^*$. Thus we have proved (i).

(ii) For $E \in \mathcal{A}^*$ define

$$\mu^*(E) = \mu(A), \ \textit{where } A, B \in \mathcal{A},\ A \subset E \subset B,\ \mu(B - A) = 0. \tag{3.1}$$

This is the only possibility for defining a measure μ^* on $\mathcal{A}^*$ that agrees with μ on $\mathcal{A}$ because $\mu(A) = \mu(B)$ whenever $A, B \in \mathcal{A}$ such that $A \subset B$ and $\mu(B - A) = 0$. To prove that μ^* is well defined, let $A_1, B_1 \in \mathcal{A}$ be another pair such that $A_1 \subset E \subset B_1$ and $\mu(B_1 - A_1) = 0$, then $A - A_1 \subset E - A_1 \subset B_1 - A_1$ and hence $\mu(A - A_1) = 0$. This implies $\mu(A) = \mu(A \cap A_1) = \mu(A_1)$. Thus the map μ^* is well defined.

We prove that μ^* is a measure. Let $E_i \in \mathcal{A}^*$ be a sequence of pairwise disjoint sets and choose sequences $A_i, B_i \in \mathcal{A}$ such that $A_i \subset E_i \subset B_i$ and $\mu(B_i - A_i) = 0$ for all i. Then the A_i are pairwise disjoint and $\mu^*(E_i) = \mu(A_i)$ for all i. Moreover, $A = \cup_i A_i,\ B = \cup_i B_i \in \mathcal{A}$, $A \subset \cup_i E_i \subset B$ and $\mu(B - A) = 0$. Hence $\mu^*(\cup_i E_i) = \mu(A) = \sum_i \mu(A_i) = \sum_i \mu^*(E_i)$. This proves (ii).

(iii) Let $E \in \mathcal{A}^*$ such that $\mu^*(E) = 0$ and let $E_0 \subset E$. Choose $A, B \in \mathcal{A}$ such that $A \subset E \subset B$ and $\mu(B - A) = 0$. Then $\mu(A) = \mu^*(E) = 0$ and hence $\mu(B) = \mu(A) + \mu(B - A) = 0$. Since $\emptyset \subset E_0 \subset E \subset B$, this implies that $E_0 \in \mathcal{A}^*$. This shows that $(X, \mathcal{A}^*, \mu^*)$ is a complete measure space.

□

A probability or probability measure is a measure μ such that $\mu(X) = 1$. In this case we usually write $(\Omega, \mathcal{F}, \mathbb{P})$ instead of $(X, \mathcal{A}, \mu)$, and $\mathcal{F}$ is called a σ-field, which is the same thing as a σ-algebra.

Exercises 2

1. If $\mu_1, \cdots, \mu_n$ are measures and $a_1, \cdots, a_n \in (0, \infty)$, then $\sum_{i=1}^n a_i \mu_i$ is a measure.

2. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{E_j\}_{j=1}^\infty \subset \mathcal{A}$.

i) Show that $\mu(\liminf E_j) \leq \liminf \mu(E_j)$.

ii) Assume that there exists $A \in \mathcal{A}$ with $\mu(A) < \infty$ such that $E_n \subset A$ for $n \in \mathbb{N}$. Show that $\mu(\limsup E_j) \geq \limsup \mu(E_j)$.

3. If $(X, \mathcal{A}, \mu)$ is a measure space and $E, F \in \mathcal{A}$, then $\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F)$.

4. Let $\mathcal{A}$ ba a σ-algebra on X. Let $\mu : \mathcal{A} \to [0,\ \infty]$ be a finite additive set function. Show that if $\{A_n\}_{n=1}^\infty$ is a family of disjoint subsets in $\mathcal{A}$ then

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

5. Let $(X, \mathcal{A}, \mu)$ be a probability space and $\{A_j\}_{j=1}^\infty \subset \mathcal{A}$ be a sequence of sets such that $\mu(A_j) = 1, \forall j$. Show that $\mu(\cap_{j=1}^\infty A_j) = 1$.

6. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{A_j\}_{j=1}^{\infty}, \{B_j\}_{j=1}^{\infty} \subset \mathcal{A}$ such that $B_j \subset A_j, \forall j$. Show that

$$\mu\left(\cup_{j=1}^{\infty} A_j\right) - \mu\left(\cup_{j=1}^{\infty} B_j\right) \leq \sum_{j=1}^{\infty}(\mu(A_j) - \mu(B_j)). \tag{3.2}$$

7. Let $\mathcal{A}$ be a σ-algebra of subsets of X and $\psi : \mathcal{A} \to [0,\ \infty]$ be a finitely additive set function. Prove that ψ is a measure if the following condition holds: If $\{A_n\}$ is a decreasing sequence of sets in $\mathcal{A}$ with $\cap_{n=1}^{\infty} A_n = \emptyset$, then $\lim_{n\to\infty} \psi(A_n) = 0$.

8. Assume that $E_1, E_2, \cdots$ are measurable subsets of the measure space $(X, \mathcal{A}, \mu)$ with the property that for $n \neq m,\ \mu(E_n \cap E_m) = 0$. Show that

$$\mu(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n).$$

9. Let $(X, \mathcal{A})$ be a measurable space and assume that $\mu : \mathcal{A} \to [0,\ \infty]$ finitely additive and σ-subadditive. Show that μ is σ-additive.

10. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $\mathcal{S} = \{A \in \mathcal{M} : \mu(A) = 0\ or\ \mu(A) = \mu(X)\}$. Prove that $\mathcal{S}$ is a σ-algebra of subsets of X.

11. Let $\mathcal{A}$ be an algebra of subsets of X, μ a nonnegative set function on $\mathcal{A}$ such that $\mu(\emptyset) = 0$, and $\mathcal{C} = \{C \in \mathcal{A} : \mu(A) = \mu(A \cap C) + \mu(A - C)\ for\ all\ A \in \mathcal{A}\}$. Prove: (a) $\mathcal{C}$ is an algebra of subsets of X. (b) The restriction of μ to $\mathcal{C}$ is finitely additive, (c) If $\{C_1, ..., C_n\} \subset \mathcal{C}$ are pairwise disjoint and $A \in \mathcal{A}$, then $\mu(\cup_{k=1}^{n}(C_k \cap A))) = \sum_{k=1}^{n} = \mu(C_k \cap A)$.

12. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and ν a positive, finitely additive set function on $\mathcal{A}$. Prove that if for any sequence $\{A_n\} \subset \mathcal{A}$ with $\mu(A_n) \to 0$ we have $\nu(A_n) \to 0$, then ν is a measure on $\mathcal{A}$.

13. Let $(X, \mathcal{A}, \mu)$ be a measure space and $T : X \to Y$ a mapping from X onto Y. Prove that $\mathcal{N} = \{B \subset Y : T^{-1}(B) \in \mathcal{A}\}$ is a σ-algebra of subsets of Y and that the set function ν on $\mathcal{N}$ given by $\nu(B) = \mu(T^{-1}(B))$ is a measure, called the measure induced by T.

14. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\mathcal{B}$ an algebra of subsets of X with $=\mathcal{M}(\mathcal{B})$ such that X is σ-finite relative to $\mathcal{B}$. Prove that $\mathcal{B}$ is dense in $\mathcal{A}$ in the sense that given $\epsilon > 0$ and $A \in \mathcal{A}$ with $\mu(A) < \infty$, there is $B \in \mathcal{B}$ such that $\mu(A\Delta B) < \epsilon$.

15. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{A_n\} \subset \mathcal{A}$ such that each A_n intersects at most one other A_m with $n \neq m$. Prove that $\mu(\cup_n A_n) \leq 2\mu(A_n) \leq 2\mu(\cup_n A_n)$.

16. Let $(X, \mathcal{A}, \mu)$ be a complete measure space. Prove: (a) If $A \cup N \in \mathcal{A}$ where $\mu(N) = 0$, then $A \in \mathcal{A}$. (b) If $A, B \subset X$ are such that $A \in \mathcal{A}$ and $\mu(A\Delta B) = 0$, then $B \in \mathcal{A}$ and $\mu(B) = \mu(A)$.

17. Let $(X, \mathcal{A}, \mu)$ be a measure space and $(X, \mathcal{S}, \mu_1)$ its completion. For $A \subset X$ define $\mu^{\star}(A) = \inf\{\mu(M) : M \in \mathcal{A}, A \subset M\}$ and $\mu_{\star}(A) = \sup\{\mu(M) : M \in \mathcal{A}, A \subset M\}$. Prove: (a) If $A \in \mathcal{S}$, then $\mu^{\star}(A) = \mu_{\star}(A)$. (b) Conversely, if $A \subset X$ is such that $\mu^{\star}(A) = \mu_{\star}(A) < \infty$, then $A \in \mathcal{S}$.

18. Let $\mathcal{A}$ be a σ-algebra of subsets of X and μ, ν measures on $\mathcal{A}$ with $\nu(A) \leq \mu(A)$ for $A \in \mathcal{A}$. Prove that there exists a measure λ on $\mathcal{A}$ such that $\mu(A) = \nu(A) + \lambda(A)$ for $A \in \mathcal{A}$.

19. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $\{A_n\} \subset \mathcal{A}$ such that $\sum_n \mu(A_n - A_{n+1}) < \infty$. Prove: (a) $\mu(\liminf A_n) = \mu(\limsup A_n) = L$. (b) $\lim \mu(A_n) = L$.

20. Let $(X, \mathcal{A}, \mu)$ be a probability measure space Given $A, B \in \mathcal{A}$, we say that $A \sim B$ iff $\mu(A \Delta B) = 0$. (a) Prove that $\sim$ is an equivalence relation on $\mathcal{A}$. (b) Consider the quotient space $\tilde{\mathcal{A}} = \mathcal{A}/\sim$ and for $[A], [B] \in \mathcal{A}$ let $d([A], [B]) = \mu(A \Delta B)$. Prove that $(\tilde{\mathcal{A}}, d)$ is a complete metric space.

21. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space so that there exists a sequence $\{E_n\} \subset \mathcal{A}$ such that $\cup_{n=1}^{\infty} E_n = X$ and $\mu(E_n) < \infty$ for every $n \in \mathbb{N}$. Show that there exists a disjoint sequence $\{F_n\} \subset \mathcal{A}$ such that $\cup_{n=1}^{\infty} F_n = X$ and $\mu(F_n) < \infty$ for every $n \in \mathbb{N}$.

22. Let $(X, \mathcal{A}, \mu)$ be a measure space and $F \in \mathcal{A}$. Show that $\mathcal{A} \ni A \to \mu(A \cap F)$ defines a measure.

4. The monotone class Theorem

This section will introduce the *monotone class, d-system and π-system* and some interesting results related which will be used later.

Definition 4.1 A monotone class is a collection of subsets $\mathcal{M}$ of X such that
(1) if $A_i \uparrow A$ and each $A_i \in \mathcal{M}$, then $A \in \mathcal{M}$;
(2) if $A_i \downarrow A$ and each $A_i \in \mathcal{M}$, then $A \in \mathcal{M}$.

The intersection of monotone classes is a monotone class, and the intersection of all monotone classes containing a given collection of sets $\mathcal{C}$ is the smallest monotone class containing that collection and is denoted by $\mathcal{M}(\mathcal{C})$.

The next theorem, the monotone class theorem, is rather technical, but very useful.

Theorem 4.2 *Suppose $\mathcal{A}_0$ is an algebra, $\mathcal{A} = \sigma(\mathcal{A}_0)$, and $\mathcal{M} \equiv \mathcal{M}(\mathcal{A}_0)$ is the smallest monotone class containing $\mathcal{A}_0$. Then $\mathcal{M} = \mathcal{A}$.*

Proof. Let $\mathcal{N}_1 = \{A \in \mathcal{M} : A^c \in \mathcal{M}\}$.

Note $\mathcal{N}_1$ is contained in $\mathcal{M}$ and contains $\mathcal{A}_0$. If $A_i \uparrow A$ and each $A_i \in \mathcal{N}_1$, then each $A_i^c \in \mathcal{M}$ and $A_i^c \downarrow A^c$. Since $\mathcal{M}$ is a monotone class, $A^c \in \mathcal{M}$, and so $A \in \mathcal{N}_1$. Similarly, if $A_i \downarrow A$ and each $A_i \in \mathcal{N}_1$, then $A \in \mathcal{N}_1$. Therefore $\mathcal{N}_1$ is a monotone class. Hence $\mathcal{N}_1 = \mathcal{M}$, and we conclude that $\mathcal{M}$ is closed under the operation of taking complements.

Let $\mathcal{N}_2 = \{A \in \mathcal{M} : A \cap B \in \mathcal{M}$ for all $B \in \mathcal{A}_0\}$.

Note that $\mathcal{N}_2$ is contained in $\mathcal{M}$ and contains $\mathcal{A}_0$ because $\mathcal{A}_0$ is an algebra. If $A_i \uparrow A$, each $A_i \in \mathcal{N}_2$ and $B \in \mathcal{A}_0$, then $A \cap B = \cup_{i=1}^{\infty}(A_i \cap B)$. Because $\mathcal{M}$ is a monotone class, $A \cap B \in \mathcal{M}$, which implies $A \in \mathcal{N}_2$. We use a similar argument when $A_i \downarrow A$.

Therefore $\mathcal{N}_2$ is a monotone class and so $\mathcal{N}_2 = \mathcal{M}$. In other words, if $B \in \mathcal{A}_0$ and $A \in \mathcal{M}$, then $A \cap B \in \mathcal{M}$.

Let $\mathcal{N}_3 = \{A \in \mathcal{M} : A \cap B \in \mathcal{M} \text{ for all } B \in \mathcal{M}\}$.

As before, $\mathcal{N}_3$ is a monotone class contained in $\mathcal{M}$ and contains $\mathcal{A}_0$ Thus $\mathcal{M}$ is a monotone class closed under the operations of taking complements and taking finite intersections. If $A_1, A_2, ...$ are elements of $\mathcal{M}$, then $B_n = A_1 \cap ... \cap A_n \in \mathcal{M}$ for each n and $B_n \downarrow \cap_{i=1}^{\infty} A_i$. Since $\mathcal{M}$ is a monotone class, we have that $\cap_{i=1}^{\infty} A_i \in \mathcal{M}$. This shows that $\mathcal{M}$ is a σ-algebra, and so $\mathcal{A} \subset \mathcal{M}$. $\square$

Definition 4.3 Let X be a set. A collection $\mathcal{D}$ of subsets of X is a d-system (or a Dynkin class) on X if

(a) $X \in \mathcal{D}$,

(b) $A - B \in \mathcal{D}$ whenever $A, B \in \mathcal{D}$ and $B \subset A$, and

(c) $\cup_{i=1}^{\infty} A_i \in \mathcal{D}$ whenever $\{A_i\}$ is an increasing sequence of sets in $\mathcal{D}$.

A collection of subsets of X is a π-system on X if it is closed under finite intersections.

Example 4.4 Suppose $\mathcal{A}$ is a σ-algebra on X. Then $\mathcal{A}$ is certainly a d-system. Furthermore, if μ and ν are finite measures on $\mathcal{A}$ such that $\mu(X) = \nu(X)$, then the collection $\mathcal{S}$ of all sets A that belong to $\mathcal{A}$ and satisfy $\mu(A) = \nu(A)$ is a d-system. The fact that such families $\mathcal{S}$ are d-systems forms the basis for many of the applications of d-systems.

The intersection of a nonempty family of d-systems on a set X is a d-system on X and that an arbitrary collection of subsets of X is included in some d-system on X, namely $\mathcal{P}(X)$, the collection of all subsets of X. Hence if $\mathcal{C}$ is an arbitrary collection of subsets of X, then the intersection of all the d-systems on X that include $\mathcal{C}$ is a d-system on X that includes $\mathcal{C}$; this intersection is the smallest such d-system and is called the d-system generated by $\mathcal{C}$ and denoted by $\mathcal{D}(\mathcal{C})$.

Theorem 4.5 *Let X be a set and let $\mathcal{C}$ be a π-system on X. Then the σ-algebra generated by $\mathcal{C}$ coincides with the d-system generated by $\mathcal{C}$.*

Proof. Let $\mathcal{D}$ be the d-system generated by $\mathcal{C}$, and, as usual, let $\sigma(\mathcal{C})$ be the σ-algebra generated by $\mathcal{C}$. Since every σ-algebra is a d-system, the σ-algebra $\sigma(\mathcal{C})$ is a d-system that includes $\mathcal{C}$, hence $\mathcal{D} \subset \sigma(\mathcal{C})$.

We prove the reverse inclusion by showing that $\mathcal{D}$ is a σ-algebra, for then $\mathcal{D}$, as a σ-algebra that includes $\mathcal{C}$, must include the σ-algebra generated by $\mathcal{C}$, namely $\sigma(\mathcal{C})$.

We begin by showing that $\mathcal{D}$ is closed under finite intersections. Define

$$\mathcal{D}_1 = \{A \in \mathcal{D} : A \cap C \in \mathcal{D} \ for \ each \ C \in \mathcal{C}\}.$$

The fact that $\mathcal{C} \subset \mathcal{D}$ implies that $X \in \mathcal{D}_1$; furthermore, the identities

$$(A - B) \cap C = (A \cap C) - (B \cap C)$$

and

$$(\cup_n A_n) \cap C = \cup_n (A_n \cap C),$$

together with the fact that $\mathcal{D}$ is a d-system, imply that $\mathcal{D}_1$ is closed under the formation of proper differences and under the formation of unions of increasing sequences of sets. Thus $\mathcal{D}_1$ is a d-system. Since $\mathcal{C}$ is closed under finite intersections and is included in $\mathcal{D}$, it is included in $\mathcal{D}_1$. Thus $\mathcal{D}_1$ is a d-system that includes $\mathcal{C}$; hence it must include $\mathcal{D}$. Therefore for any $A \in \mathcal{D}$ and any $C \in \mathcal{C}$ we have $A \cap C \in \mathcal{D}$.

Next we define

$$\mathcal{D}_2 = \{B \in \mathcal{D} : A \cap B \in \mathcal{D} \textit{ for each } A \in \mathcal{D}\}.$$

The previous step of this proof shows that $\mathcal{C} \subset \mathcal{D}_2$, and a straightforward modification of the argument in the previous step shows that $\mathcal{D}_2$ is a d-system. It follows that $\mathcal{D} \subset \mathcal{D}_2$, that is, $\mathcal{D}$ is closed under finite intersections.

Parts (a) and (b) of the definition of a d-system imply that $X \in \mathcal{D}$ and that $\mathcal{D}$ is closed under complementation. As we have just seen, $\mathcal{D}$ is also closed under the formation of finite intersections, and so it is an algebra. Finally $\mathcal{D}$, as a d-system, is closed under unions of increasing sequences of sets, it must be a σ-algebra. The proof is complete. □

Corollary 4.6. *Let $(X, \mathcal{A})$ be a measurable space, and let $\mathcal{C}$ be a π-system on X such that $\mathcal{A} = \sigma(\mathcal{C})$. If μ and ν are finite measures on $(X, \mathcal{A})$ that satisfy $\mu(X) = \nu(X)$ and satisfy $\mu(A) = \nu(A)$ for each $A \in \mathcal{C}$. Then $\mu = \nu$.*

Proof. Let

$$\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\};$$

then $\mathcal{D}$ is a d-system. Since $\mathcal{C}$ is a π-system and is included in $\mathcal{D}$, it follows from Theorem 4.5 that $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{A}$. Thus $\mu(A) = \nu(A)$ holds for each A in $\mathcal{A}$, and the proof is complete. □

Corollary 4.7. *Let $(X, \mathcal{A})$ be a measurable space, and let $\mathcal{C}$ be a π-system on X such that $\mathcal{A} = \sigma(\mathcal{C})$. If μ and ν are measures on $(X, \mathcal{A})$ that agree on $\mathcal{C}$, and if there is an increasing sequence $\{C_n\}$ of sets that belong to $\mathcal{C}$, have finite measure under μ and ν, and satisfy $\cup_{n=1}^{\infty} C_n = X$, then $\mu = \nu$.*

Proof. For each positive integer n define measures μ_n and ν_n on $\mathcal{A}$ by $\mu_n(A) = \mu(A \cap C_n)$ and $\nu_n(A) = \nu(A \cap C_n)$. Corollary 4.6 implies that for each n we have $\mu_n = \nu_n$. Since

$$\mu(A) = \lim_{n \to \infty} \mu_n(A) = \lim_{n \to \infty} \nu_n(A) = \nu(A)$$

holds for each A in $\mathcal{A}$, μ and ν must be equal. □

Exercises 3

1. Find an example of a set X and a monotone class $\mathcal{M}$ consisting of subsets of X such that $\emptyset, X \in \mathcal{M}$, but $\mathcal{M}$ is not a σ-algebra.

2. Let X be a set and let $(A_i)_{i \in I}$ be a partition of X, i.e., A_i is a nonempty subset of X for each $i \in I, A_i \cap A_j = \emptyset$; for $i \neq j$, and $X = \cup_{i=1}^{\infty} A_i$. Then

$$\mathcal{A} = \{A_J := \cup_{j \in J} A_j | J \subset I\}$$

is a σ-algebra.

3. Prove that if $\mathcal{A}$ is an algebra of subsets of X and μ, ν are finite measures on $\sigma(\mathcal{A})$ such that $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$, then $\mu = \nu$.

4. Let $\mathcal{A}$ be an algebra of subsets of X and μ, ν measures on $\sigma(\mathcal{A})$ that are σ-finite relative to $\mathcal{A}$, i.e., $X = \cup_{n=1}^{\infty} A_n$ with the A_n pairwise disjoint sets in $\mathcal{A}$ and $\mu(A_n) = \nu(A_n) < \infty$ for all n. Prove that if $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$, then $\mu = \nu$.

5. Let $\mathcal{F}$ be a π-system on X such that $X \in \mathcal{F}$ and if $A \in \mathcal{F}$, then $A^c = \cup_{k=1}^{n} A_k$ with $A_k \in \mathcal{F}$. Prove that $\mathcal{A} = \{A \subset X : A = \cup_{k=1}^{n} A_k, A_k \in \mathcal{F}\} = \mathcal{A}(\mathcal{F})$, the algebra of subsets of X generated by $\mathcal{F}$.

6. Let $\mathcal{C}$ be a collection of subsets of X and $\mathcal{M}$ is a monotone class of a collection of subsets of X. Show that $\mathcal{N} = \{A \in \mathcal{M} : A \cap C \in \mathcal{M} \ for\ all\ C \in \mathcal{C}\}$ is a monotone class. Show that given a collection $\mathcal{C}$ of subsets of X, there is a smallest monotone class of subsets of X that contains $\mathcal{C}$. If a monotone class $\mathcal{M}$ is an algebra of subsets of X, then $\mathcal{M}$ is a σ-algebra.

7. Let $(X, \mathcal{A}, \mu)$ be a probability measure space and let $\mathcal{F}, \mathcal{F}_1 \subset \mathcal{A}$ be π-systems in X. Prove that $\mu(B \cap C) = \mu(B)\mu(C)$ for all $B \in \sigma(\mathcal{F}), C \in \sigma(\mathcal{F}_1)$ iff $\mu(B \cap C) = \mu(B)\mu(C)$ for all $B \in \mathcal{F}, C \in \mathcal{F}_1$.

8. Show that a d-system is a σ-algebra if, and only if, it is closed with respect to finite intersections.

5. Construction of outer measures

The purpose of this section is to give a method for constructing measures. This is a complicated procedure, and involves the concept of outer measure, which we introduce in the first part of this section. The most important example is the one-dimensional Lebesgue measure.

The methods used to construct measures via outer measures have other applications besides the construction of Lebesgue measure. The Carathéodory extension theorem is a tool in constructing measures.

Definition 5.1 Let X be a set. An outer measure is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ satisfying

(1) $\mu^*(\emptyset) = 0$;
(2) if $A \subset B$, then $\mu^*(A) \le \mu^*(B)$;
(3) $\mu^*(\cup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu^*(A_i)$ whenever $A_1, A_2, ...$ are subsets of X.

Example 5.2 (a) Let X be an arbitrary set, and define μ^* on $\mathcal{P}(X)$ by $\mu^*(A) = 0$ if $A = \emptyset$ and $\mu^*(A) = 1$ otherwise. Then μ^* is an outer measure.

(b) Let X be an arbitrary set, and define μ^* on $\mathcal{P}(X)$ by $\mu^*(A) = 0$ if A is countable, and $\mu^*(A) = 1$ if A is uncountable. Then μ^* is an outer measure.

(c) Let X be an infinite set, and define μ^* on $\mathcal{P}(X)$ by $\mu^*(A) = 0$ if A is finite, and $\mu^*(A) = 1$ if A is infinite. Then μ^* fails to be countably sub-additive and so is not an outer measure.

A set N is a null set with respect to μ^* if $\mu^*(N) = 0$.

The following result is a common way to generate outer measures.

Proposition 5.3 *Suppose $\mathcal{C}$ is a collection of subsets of X such that $\emptyset \in \mathcal{C}$ and there exist $D_1, D_2, \ldots$ in $\mathcal{C}$ such that $X = \cup_{i=1}^\infty D_i$. Suppose $l : \mathcal{C} \to [0, \infty]$ with $l(\emptyset) = 0$. Define*

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^\infty l(A_i) : A_i \in \mathcal{C} \text{ for each } i \text{ and } E \subset \bigcup_{i=1}^\infty A_i \right\}. \tag{5.1}$$

Then μ^* is an outer measure.

Proof. (1) and (2) of the definition of outer measure are obvious. To prove (3), let $A_1, A_2, \ldots$ be subsets of X and let $\epsilon > 0$. For each i there exist $C_{i1}, C_{i2}, \ldots \in \mathcal{C}$ such that $A_i \subset \cup_{j=1}^\infty C_{ij}$ and $\sum_{j=1}^\infty l(C_{ij}) \le \mu^*(A_i) + \epsilon/2^i$. Then $\cup_{i=1}^\infty A_i \subset \cup_i \cup_j C_{ij}$ and

$$\begin{aligned} \mu^*(\cup_{i=1}^\infty A_i) &\le \sum_{i,j} l(C_{ij}) = \sum_i \sum_j l(C_{ij}) \\ &\le \sum_{i=1}^\infty \mu^*(A_i) + \sum_{i=1}^\infty \frac{\epsilon}{2^i} = \sum_{i=1}^\infty \mu^*(A_i) + \epsilon. \end{aligned}$$

Since ϵ is arbitrary, $\mu^*(\cup_{i=1}^\infty A_i) \le \sum_{i=1}^\infty \mu^*(A_i)$. □

Definition 5.4 Let μ^* be an outer measure. A set $A \subset X$ is μ^*-measurable if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{5.2}$$

for all $E \subset X$.

The following is sometimes known as the *Carathéodory theorem*.

Theorem 5.5 *If μ^* is an outer measure on X, then the collection $\mathcal{A}$ of μ^*-measurable sets is a σ-algebra. If μ is the restriction of μ^* to $\mathcal{A}$, then μ is a complete measure.*

Proof. By Definition 5.1,

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

for all $E \subset X$. Thus to check (5.2) it is enough to show

$$\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c) \tag{5.3}$$

This is trivial in the case $\mu^*(E) = \infty$.

Step 1. We show $\mathcal{A}$ is an algebra. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ by symmetry and the definition of $\mathcal{A}$. Suppose $A, B \in \mathcal{A}$ and $E \subset X$. Then

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) \\ &\quad + \mu^*(E \cap A^c \cap B^c). \end{aligned} \tag{5.4}$$

Observing

$$\mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B^c) \ge \mu^*(E \cap B^c) \tag{5.5}$$

and

$$B^c \cup (A^c \cap B) = B^c \cup A^c = (A \cap B)^c, \tag{5.6}$$

we obtain that

$$\begin{aligned}\mu^*(E) &\geq \mu^*(E\cap A\cap B)+\mu^*(E\cap A^c\cap B)+\mu^*(E\cap B^c)\\ &\geq \mu^*(E\cap A\cap B)+\mu^*((E\cap A^c\cap B)\cup(E\cap B^c))\\ &= \mu^*(E\cap A\cap B)+\mu^*(E\cap(A\cap B)^c),\end{aligned}$$

which shows that $A\cap B\in\mathcal{A}$. Therefore $\mathcal{A}$ is an algebra.

Step 2. Next we show $\mathcal{A}$ is a σ-algebra. Let A_i be pairwise disjoint sets in $\mathcal{A}$, let $B_n=\cup_{i=1}^n A_i$ and $B=\cup_{i=1}^\infty A_i$. If $E\in\mathcal{P}(X)$,

$$\begin{aligned}\mu^*(E\cap B_n) &= \mu^*(E\cap B_n\cap A_n)+\mu^*(E\cap B_n\cap A_n^c)\\ &= \mu^*(E\cap A_n)+\mu^*(E\cap B_{n-1})\\ &= \cdots=\sum_{i=1}^n\mu^*(E\cap A_i).\end{aligned}$$

Since $B_n\in\mathcal{A}$, then

$$\begin{aligned}\mu^*(E) &= \mu^*(E\cap B_n)+\mu^*(E\cap B_n^c)=\sum_{i=1}^n\mu^*(E\cap A_i)+\mu^*(E\cap B_n^c)\\ &\geq \sum_{i=1}^n\mu^*(E\cap A_i)+\mu^*(E\cap B^c).\end{aligned}\tag{5.7}$$

Letting $n\to\infty$ and recalling that μ^* is an outer measure,

$$\begin{aligned}\mu^*(E) &\geq \sum_{i=1}^\infty\mu^*(E\cap A_i)+\mu^*(E\cap B^c)\geq\mu^*(\bigcup_{i=1}^\infty(E\cap A_i))+\mu^*(E\cap B^c)\\ &= \mu^*(E\cap B)+\mu^*(E\cap B^c)\geq\mu^*(E).\end{aligned}\tag{5.8}$$

This shows $B\in\mathcal{A}$ and so $\mathcal{A}$ is a σ-algebra.

Step 3. We now show μ^* restricted to $\mathcal{A}$ is a measure. The only way (5.8) can hold is if all the inequalities there are actually equalities, and in particular,

$$\mu^*(E)=\sum_{i=1}^\infty\mu^*(E\cap A_i)+\mu^*(E\cap B^c).$$

Taking $E=B$, we obtain

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

This shows that μ^* is countably additive on $\mathcal{A}$.

Step 4. Finally, if $\mu^*(A)=0$ and $E\subset X$, then

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

which shows that $\mathcal{A}$ contains all the sets of μ^*-measure zero. □

Example 5.6 Let $X=\mathbb{R}$ and let $\mathcal{C}$ be the collection of intervals of the form $(a,b]$. Let $l(I)=b-a$ if $I=(a,b]$. Define μ^* by (5.1). Proposition 5.3 shows that μ^* is an outer measure and Theorem 5.5 tells us that μ^* is a complete measure on the σ-algebra $\mathcal{L}$, the collection of subsets of $\mathbb{R}$ which are μ^* measurable. This measure, denoted by m, is known as Lebesgue measure. The σ-algebra $\mathcal{L}$ is called the Lebesgue σ-algebra. There are subsets in $\mathbb{R}$ which are non-measurable with respect to μ^* and so μ^* is not a measure on $\mathcal{P}(\mathbb{R})$.

We introduce the Carathéodory extension theorem now. This theorem gives a tool for constructing measures in a variety of contexts.

Let $\mathcal{A}_0$ be an algebra. Saying $\mu : \mathcal{A}_0 \to [0, \infty]$ is a measure on $\mathcal{A}_0$ means the following: $\mu(\emptyset) = 0$ and if $A_1, A_2, ...$ are pairwise disjoint elements of $\mathcal{A}_0$ and also $\cup_{i=1}^{\infty} A_i \in \mathcal{A}_0$, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$. Sometimes one calls a measure on an algebra a *premeasure* .

Theorem 5.7 (Caratheodory-Hahn) *Suppose $\mathcal{A}_0$ is an algebra on X and $l : \mathcal{A}_0 \to [0, \infty]$ is a measure. For $E \in \mathcal{P}(X)$, define*

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} l(A_i) : A_i \in \mathcal{A}_0 \ for\ each\ i\ and\ E \subset \bigcup_{i=1}^{\infty} A_i \right\}. \tag{5.9}$$

Then

(1) μ^* is an outer measure;

(2) $\mu^*(A) = l(A)$ if $A \in \mathcal{A}_0$;

(3) every set in $\mathcal{A}_0$ is μ^*-measurable;

(4) if l is σ-finite, then μ^* is a measure on $\sigma(\mathcal{A}_0)$ which is the unique extension of l to $\sigma(\mathcal{A}_0)$.

Proof. (1) is Proposition 5.3. We turn to (2). Suppose $E \in \mathcal{A}_0$. We know $\mu^*(E) \le l(E)$ since we can take $A_1 = E$ and $A_2, A_3, ...$ empty in the definition of μ^*. If $E \subset \cup_{i=1}^{\infty} A_i$ with $A_i \in \mathcal{A}_0$, let $B_1 = E \cap A_1$ and $B_n = E \cap (A_n - (\cup_{i=1}^{n-1} A_i)), n \ge 2$. Then the B_n are pairwise disjoint, they are each in $\mathcal{A}_0$, and their union is E. Therefore

$$l(E) = \sum_{i=1}^{\infty} l(B_i) \le \sum_{i=1}^{\infty} l(A_i).$$

Taking the infimum over all such sequences $\{A_i\}$, shows that

$$l(E) \le \mu^*(E).$$

Next we prove (3). Let $A \in \mathcal{A}_0$, $\epsilon > 0$ and $E \subset X$. Pick $B_1, B_2, ... \in \mathcal{A}_0$ such that $E \subset \bigcup_{i=1}^{\infty} B_i$ and $\sum_{i=1}^{\infty} l(B_i) \le \mu^*(E) + \epsilon$. Then

$$\begin{aligned} \mu^*(E) + \epsilon \ge \sum_{i=1}^{\infty} l(B_i) &= \sum_{i=1}^{\infty} l(B_i \cap A) + \sum_{i=1}^{\infty} l(B_i \cap A^c) \\ &\ge \mu^*(E \cap A) + \mu^*(E \cap A^c). \end{aligned}$$

Since ϵ is arbitrary, $\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Thus, A is μ^*-measurable.

Finally, we prove (4). Suppose we have two extensions to $\sigma(\mathcal{A}_0)$, the σ-algebra generated by $\mathcal{A}_0$. One is μ^* and let the other extension be called ν. We will show that if $E \in \sigma(\mathcal{A}_0)$, then $\mu^*(E) = \nu(E)$. Let us first assume that μ^* is a finite measure. The μ^*-measurable sets form a σ-algebra containing $\mathcal{A}_0$. Because

$E \in \sigma(\mathcal{A}_0)$, E must be μ^*-measurable and

$$\mu^*(E) = \inf \left\{ \sum_{i=1}^{\infty} l(A_i) : A_i \in \mathcal{A}_0 \ for\ each\ i\ and\ E \subset \bigcup_{i=1}^{\infty} A_i \right\}. \tag{5.10}$$

But $l = \nu$ on $\mathcal{A}_0$, so $\sum_{i=1}^{\infty} l(A_i) = \sum_{i=1}^{\infty} \nu(A_i)$. Therefore if $E \subset \bigcup_{i=1}^{\infty} A_i$ with each $A_i \in \mathcal{A}_0$, then

$$\nu(E) \le \sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} l(A_i)$$

which implies

$$\nu(E) \le \mu^*(E). \tag{5.11}$$

Thus we also have

$$\nu(E^c) \le \mu^*(E^c). \tag{5.12}$$

Adding the above two inequalities, we have

$$\nu(X) = \nu(E) + \nu(E^c) \le \mu^*(E) + \mu^*(E^c) = \mu^*(X) = l(X) = \nu(X). \tag{5.13}$$

Therefore $\nu = \mu^*$ on $\sigma(\mathcal{A}_0)$.

It remains to consider the case when l is σ-finite. Write $X = \bigcup_{i=1}^{\infty} X_i$, where $X_i \uparrow X$ and $l(X_i) < \infty$. By the preceding paragraph we have unique extension for the measure l_i defined by $l_i(A) = l(A \cap X_i)$ to $\sigma(\mathcal{A}_0)$. If μ and ν are two extensions of l, then $\mu_i, \nu_i : \sigma(\mathcal{A}_0) \to [0, \infty]$ given by $\mu_i(B) = \mu(B \cap X_i)$, $\nu_i(B) = \nu(B \cap X_i)$, $B \in \sigma(\mathcal{A}_0)$, are extensions of l_i to $\sigma(\mathcal{A}_0)$ since they are equal on $\mathcal{A}_0$. By the uniqueness above, we know that $\mu_i = \nu_i$. Hence, for any $A \in \sigma(\mathcal{A}_0)$, we have

$$\mu(A) = \lim_{i \to \infty} \mu(A \cap X_i) = \lim_{i \to \infty} \nu(A \cap X_i) = \nu(A).$$

□

Exercises 4

In exercises 1-5, we assume that μ^* is an outer measure on some set X.

1. Let E be a measurable subset of X. Show that for every subset A of X the following equality holds:

$$\mu^*(E \cup A) + \mu^*(E \cap A) = \mu^*(E) + \mu^*(A).$$

2. If A is a nonmeasurable subset of X and E is a measurable set such that $A \subset E$, then show that $\mu^*(E - A) > 0$.

3. Let $\{A_n\}$ be a sequence of subsets of X. Assume that there exists a disjoint sequence $\{B_n\}$ of measurable sets such that $A_n \subset B_n$ holds for each n. Show that

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

4. Show that a subset E of X is measurable if and only if for each $\epsilon > 0$ there exists a measurable set F such that $F \subset E$ and $\mu^*(E - F) < \epsilon$.

5. Assume that a subset E of X has the property that for each $\epsilon > 0$, there exists a measurable set F such that $\mu^*(E \Delta F) < \epsilon$. Show that E is a measurable set.

6. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be the outer measure induced by μ. Show that for all $A \subset X$ there exists $B \in \mathcal{A}$ containing A with $\mu(B) = \mu^*(A)$.

6. Measurable functions

In analogy to continuous maps between topological spaces, one can define measurable maps between measurable spaces as those maps under which pre-images of measurable sets are again measurable.

For any map $f : X \to Y$ between two sets X and Y and any subset $B \subset Y$, we have

$$f^{-1}(Y) = X, \; f^{-1}(\emptyset) = \emptyset, \tag{6.1}$$

$$f^{-1}(Y - B) = X - f^{-1}(B) \tag{6.2}$$

and

$$f^{-1}(\cup_{i \in I} B_i) = \cup_{i \in I} f^{-1}(B_i), \; f^{-1}(\cap_{i \in I} B_i) = \cap_{i \in I} f^{-1}(B_i) \tag{6.3}$$

for every collection of subsets $B_i \subset Y$, indexed by a set I.

Consider the set $\overline{\mathbb{R}}:=[-\infty, \infty] := \mathbb{R} \cup \{-\infty, \infty\}$

For $a, b \in \mathbb{R}$ define

$$(a, \infty] := (a, \infty) \cup \{\infty\}, \; [-\infty, b) := (-\infty, b) \cup \{-\infty\}.$$

Call a subset $U \subset \overline{\mathbb{R}}$ open if it is a countable union of open intervals in $\mathbb{R}$ and sets of the form $(a, \infty]$ or $[-\infty, b)$ for $a, b \in \mathbb{R}$. The set of open subsets of $\overline{\mathbb{R}}$ satisfies the axioms of a topology. This is called the standard topology on $\overline{\mathbb{R}}$.

Definition 6.1 (measurable function). (i) Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A map $f : X \to Y$ is called *measurable* if

$$B \in \mathcal{B} \Rightarrow f^{-1}(B) \in \mathcal{A}. \tag{6.4}$$

(ii) Let $(X, \mathcal{A})$ be a measurable space. A function $f : X \to \overline{\mathbb{R}}$ is *measurable* if it is measurable with respect to the Borel σ-algebra on $\overline{\mathbb{R}}$ associated to the standard topology as defined above.

(iii) Let X and Y be topological spaces. A map $f : X \to Y$ is called *Borel measurable* if the pre-image of every Borel measurable subset of Y under f is a Borel measurable subset of X.

Example 6.2 Let X be a set. The characteristic function of a subset $A \subset X$ is the function

$$\chi_A : X \to \mathbb{R}$$

defined by

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases} \tag{6.5}$$

Now assume $(X, \mathcal{A})$ is a measurable space, consider the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$, and let $A \subset X$ be any subset. If $B \subset \mathbb{R}$, then

$$\chi_A^{-1}(B) = \begin{cases} A, & if\ 0 \notin B,\ 1 \in B \\ A^c, & if\ 0 \in B, 1 \notin B \\ X, & if\ 0 \in B,\ 1 \in B \\ \emptyset, & if\ 0 \notin B,\ 1 \notin B. \end{cases} \tag{6.6}$$

Thus χ_A is a measurable function if and only if A is a measurable set.

Theorem 6.3 *Let $(X, \mathcal{A}), (Y, \mathcal{B}), (Z, \mathcal{C})$ be measurable spaces.*

(i) The identity map $\mathrm{id}_X : X \to X$ is measurable.

(ii) If $f : X \to Y$ and $g : Y \to Z$ are measurable maps, then so is the composition $g \circ f : X \to Z$.

(iii) Let $f : X \to Y$ be any map. Then the set

$$f_*\mathcal{A} := \{B \subset Y | f^{-1}(B) \in \mathcal{A}\} \tag{6.7}$$

is a σ-algebra on Y, called the pushforward of $\mathcal{A}$ under f.

(iv) A map $f : X \to Y$ is measurable if and only if $\mathcal{B} \subset f_*\mathcal{A}$.

Proof. Parts (i) and (ii) follow directly from the definitions. That $f_*\mathcal{A}$ defined by (6.7) is a σ-algebra follows from (6.1)-(6.3) and Definition 2.2. This proves part (iii). Moreover by Definition 6.1, f is measurable if and only if $f^{-1}(B) \in \mathcal{A}$ for every $B \in \mathcal{B}$, and this means that $\mathcal{B} \subset f_*\mathcal{A}$. This establishes (iv) and completes the proof of Theorem 6.3. $\square$

Theorem 6.4 (measurable and continuous maps). *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Assume that $\mathcal{B}$ is the Borel σ-algebra of a topology on Y.*

(i) A map $f : X \to Y$ is measurable if and only if for every open subset $U \subset Y$, $f^{-1}(U) \in \mathcal{A}$.

(ii) Assume that $\mathcal{A} = \mathcal{B}(X)$ is the Borel σ-algebra of a topology on X. Then every continuous map $f : X \to Y$ is (Borel) measurable.

Proof. Let $\mathcal{O}$ be the collection of open subsets of Y.

(i) Assuming that for every open subset $U \subset Y$, $f^{-1}(U) \in \mathcal{A}$, we have $\mathcal{O} \subset f_*\mathcal{A}$. Since $f_*\mathcal{A}$ a σ-algebra on Y and the Borel σ-algebra $\mathcal{B}$ is the smallest σ-algebra on Y containing the collection of open sets of Y, it follows that $\mathcal{B} \subset f_*\mathcal{A}$. Thus f is measurable. The other part of (i) is trivial.

(ii) If $f : X \to Y$ is continuous, then for any $U \in \mathcal{O}$, $f^{-1}(U)$ is an open subset of X and so belongs to $\mathcal{A}$. It follows from (i) that f is measurable. $\square$

Theorem 6.5 (characterization of measurable functions). *Let $(X, \mathcal{A})$ be a measurable space and let $f : X \to \overline{\mathbb{R}}$ be any function. Then the following are equivalent:*

(i) f is measurable;

(ii) $f^{-1}((a, \infty])$ is a measurable subset of X for every $a \in \mathbb{R}$;

(iii) $f^{-1}([a, \infty])$ is a measurable subset of X for every $a \in \mathbb{R}$;

(iv) $f^{-1}([-\infty, b))$ is a measurable subset of X for every $b \in \mathbb{R}$;

(v) $f^{-1}([-\infty, b])$ is a measurable subset of X for every $b \in \mathbb{R}$.

Proof. That (i) implies (ii), (iii), (iv), and (v) follows directly from the definitions. We prove that (ii) implies (i). Observe that

$$f_*\mathcal{A} = \{B \subset \overline{\mathbb{R}} | f^{-1}(B) \in \mathcal{A}\} \subset \mathcal{P}(\overline{\mathbb{R}}) \tag{6.8}$$

is a σ-algebra on $\overline{\mathbb{R}}$ by (iii) in Theorem 6.3. Let $f : X \to \overline{\mathbb{R}}$ be such that $f^{-1}((a, \infty]) \in \mathcal{A}$ for every $a \in \mathbb{R}$; then $(a, \infty] \in f_*\mathcal{A}$ for every $a \in \mathbb{R}$. Therefore $[-\infty, b] = \overline{\mathbb{R}} - (b, \infty] \in f_*\mathcal{A}$ for every $b \in \mathbb{R}$ and so

$$[-\infty, b) = \cup_{n=1}^{\infty} \left[-\infty, b - \frac{1}{n}\right] \in f_*(\mathcal{A}).$$

Thus we have

$$(a, b) = [-\infty, b) \cap (a, \infty] \in f_*(\mathcal{A})$$

for every pair of real numbers $a < b$. Since every open subset of $\overline{\mathbb{R}}$ is a countable union of sets of the form $(a, b), (a, \infty], [-\infty, b)$, it follows that every open subset of $\overline{\mathbb{R}}$ is an element of $f_*\mathcal{A}$. Hence f is measurable. This shows that (ii) implies (i). That either of the conditions (iii), (iv), and (v) also implies (i) can be shown by a similar argument. □

Theorem 6.6 (vector valued measurable functions). *Let $(X, \mathcal{A})$ be a measurable space and let $f = (f_1, \cdots, f_n) : X \to \mathbb{R}^n$ be a function. Then f is measurable if and only if $f_i : X \to \mathbb{R}$ is measurable for each i.*

Proof. For $i = 1, ..., n$ define the projection $\pi_i : \mathbb{R}^n \to \mathbb{R}$ by $\pi_i(x_1, ..., x_n) = x_i$. Since π_i is continuous, it follows that if f is measurable, so is $f_i = \pi_i \circ f$ for all i. Conversely, suppose that f_i is measurable for $i = 1, ..., n$. Let $a_i < b_i$ for $i = 1, ..., n$ and define

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

Then

$$f^{-1}(Q(a; b)) = \cap_{i=1}^{n} f_i^{-1}((a_i, b_i)) \in \mathcal{A}.$$

Now every open subset of $\mathbb{R}^n$ can be expressed as a countable union of sets of the form $Q(a; b)$. It follows that $f^{-1}(U) \in \mathcal{A}$ for every open set $U \subset \mathbb{R}^n$, and so f is measurable. □

Lemma 6.7. *Let $(X, \mathcal{A})$ be a measurable space and let $u, v : X \to \mathbb{R}$ be measurable functions. If $\phi : \mathbb{R}^2 \to \mathbb{R}$ is continuous, then the function $h : X \to \mathbb{R}$ defined by*

$$h(x) = \phi(u(x), v(x)), \;\; x \in X$$

is measurable.

Proof. The map $f := (u, v) : X \to \mathbb{R}^2$ is measurable (with respect to the Borel σ-algebra on $\mathbb{R}^2$) and the map $\phi : \mathbb{R}^2 \to \mathbb{R}$ is Borel measurable. Hence the composition $h = \phi \circ f : X \to \mathbb{R}$ is measurable.

□

Theorem 6.8 (properties of measurable functions). *Let $(X, \mathcal{A})$ be a measurable space.*

(i) If $f, g : X \to \mathbb{R}$ are measurable functions, then so are the functions

$$f+g; \quad fg; \quad \max\{f,g\}; \quad \min\{f,g\}; \quad |f|.$$

(ii) Let $f_k : X \to \overline{\mathbb{R}}$, $k = 1, 2, \ldots$ be a sequence of measurable functions. Then the following functions from X to $\overline{\mathbb{R}}$ are measurable:

$$\inf f_k, \quad \sup f_k, \quad \limsup_k f_k, \quad \liminf_k f_k.$$

Proof. (i) The functions $\phi : \mathbb{R}^2 \to \mathbb{R}$ defined by $\phi(s,t) = s+t$, $\phi(s,t) = st$, $\phi(s,t) = \max\{s,t\}$, $\phi(s,t) = \min\{s,t\}$, $\phi(s,t) = |s|$ are all continuous. Hence assertion (i) follows from Lemma 6.7.

ii) Define

$$g := \sup_k f_k : X \to \overline{\mathbb{R}}$$

and let $a \in \mathbb{R}$. Then the set

$$\begin{aligned} g^{-1}((a,\infty]) &= \{x \in X : \sup_i f_i(x) > a\} = \{x \in X : \exists\, i \text{ such that } f_i(x) > a\} \\ &= \bigcup_{i=1}^{\infty} \{x \in X | f_i(x) > a\} = \bigcup_{i=1}^{\infty} f_i^{-1}((a,\infty]). \end{aligned}$$

is measurable. Hence it follows from Theorem 6.5 that g is measurable. It also follows from (i) that $-f_k$ is measurable, hence so is $\sup(-f_k)$ by what we have just proved, and then so is the function $\inf f_k = -\sup(-f_k)$. With this understood, we conclude that the functions

$$\limsup_k f_k = \inf_{l\in\mathbb{N}} \sup_{k\geq l} f_k; \; \liminf_k f_k = \sup_{l\in\mathbb{N}} \inf_{k\geq l} f_k$$

are also measurable. □

Proposition 6.9 *If $f : \mathbb{R} \to \mathbb{R}$ is monotone, then f is Borel measurable.*

Proof. Let us suppose f is increasing, for otherwise we look at $-f$. Given $a \in \mathbb{R}$, let $x_0 = \sup\{x : f(x) \leq a\}$. If $f(x_0) \leq a$, then $\{x : f(x) > a\} = (x_0, \infty)$. If $f(x_0) > a$, then $\{x : f(x) > a\} = [x_0, \infty)$. In either case $\{x : f(x) > a\}$ is a Borel set. □

Exercises 5

1. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f_n : X \to \mathbb{R}$ be any sequence of measurable functions. Show that

$$\begin{aligned} E : &= \{x \in X, \{f_n(x)\} \textit{ is a Cauchy sequence}\} \\ &= \cap_{k=1}^{\infty} \cup_{n_0=1}^{\infty} \cap_{n,m\geq n_0} \{x \in X, |f_n(x) - f_m(x)| < \frac{1}{2^k}\} \end{aligned}$$

and so E is measurable.

2. A measure space $(X, \mathcal{A}, \mu)$ is complete iff given f and g on X with f measurable and $g = f$ μ a.e., g is measurable.

3. Let $\mathcal{A}$ be a σ-algebra of subsets of X and f a function on X such that $|f|$ is measurable. Find an additional condition that ensures that f is measurable.

4. Given two measurable spaces $(X, \mathcal{A})$ and $(X, \mathcal{B})$. Let f be a measurable mapping of X into Y. Let μ be a measure on $\mathcal{A}$. Show that the set function defined by $\nu = \mu \circ f^{-1}$ on $\mathcal{B}$, that is, $\nu(B) = \mu(f^{-1}(B))$ for $B \in \mathcal{B}$ is a measure on $\mathcal{B}$. This measure ν is called the image measure or push forward of μ under f and is denoted by $f(\mu)$

5. Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \overline{\mathbb{R}}$ be measurable. The distribution function F of f, relative to μ, is defined by

$$F(t) = \mu(\{f > t\}).$$

(i) Show that F is nonincreasing and satisfies

$$\lim_{t\to-\infty} F(t) = \lim_{n\to-\infty} F(n) = \lim_{n\to-\infty} \mu(\{f > n\}) = \mu(\{f > -\infty\}).$$

(ii) Show that for any $t_0 \in \mathbb{R}$, we have $\lim_{t\to t^+} F(t) = F(t_0)$, that is, F is right continuous.

(iii) Show that for any $t_0 \in \mathbb{R}$, we have $\lim_{t\to t^-} F(t) = \mu(\{f \geq t_0\})$, that is, F has left limits.

6. Let X be a set, let $(X_i, \mathcal{A}_i), i \in \mathcal{I}$, be arbitrarily many measurable spaces and let $T_i : X \to X_i$ be a family of maps.

(i) Show that for every $i \in \mathcal{I}$ the smallest σ-algebra in X that makes T_i measurable is given by $T_i^{-1}(\mathcal{A}_i) = \{T_i^{-1}(A) : A \in \mathcal{A}_i\}$.

(ii) Show that $\sigma(\cup_{i\in\mathcal{I}}\mathcal{A}_i)$ is the smallest σ-algebra in X that makes all $T_i, i \in \mathcal{I}$, simultaneously measurable.

7. Let $T : X \to Y$ be a map. Show that $T^{-1}(\sigma(G)) = \sigma(T^{-1}(G))$ holds for any families G of subsets of Y.

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

(i) Let $f, g : X \to \mathbb{R}$ be measurable functions. Show that for every $A \in \mathcal{A}$ the function $h(x) := f(x),\ if\ x \in A,\ and\ h(x) := g(x),\ if\ x \notin A$, is measurable.

(ii) Let $\{f_n\}$ be a sequence of measurable functions and let $\{A_n\} \subset \mathcal{A}$ such that $\cup_n A_n = X$. Suppose that $f_n|_{A_n\cap A_k} = f_k|_{A_n\cap A_k}$ for all $k, n \in \mathbb{N}$ and set $f(x) := f_n(x)$ if $x \in A_n$. Show that $f : X \to \mathbb{R}$ is measurable.

9. Show that any left- or right-continuous function $u : \mathbb{R} \to \mathbb{R}$ is measurable

10. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous, 1-1, and onto. Prove that f maps Borel sets onto Borel sets.

7. Convergence of measurable functions

In this section, we continue on studying measurable functions, but now with an emphasis on results that depend upon measures. The highlight of this section is the proof of Egorov's Theorem.

Egorov's Theorem *Suppose $(X, \mathcal{A}, \mu)$ is a measure space with $\mu(X) < \infty$. Suppose $f_1, f_2, ...$ is a sequence of measurable functions from X to $\mathbb{R}$ that converges* *pointwise on X to a measurable function $f : X \to \mathbb{R}$. Then for every $\epsilon > 0$, there exists a set $E \in \mathcal{A}$ such that $\mu(X - E) < \epsilon$ and $f_n \to f$ uniformly on E.*

Proof. Suppose $\epsilon > 0$ and $n \in \mathbb{N}$. The definition of pointwise convergence implies that

$$\bigcup_{m=1}^{\infty} \bigcap_{k=m}^{\infty} \left\{ x \in X : |f_k(x) - f(x)| < \frac{1}{n} \right\} = X. \tag{7.1}$$

For $m \in \mathbb{N}$, let

$$A_{m,n} = \bigcap_{k=m}^{\infty} \left\{ x \in X : |f_k(x) - f(x)| < \frac{1}{n} \right\}. \tag{7.2}$$

Then $A_{1,n} \subset A_{2,n} \subset$ is an increasing sequence of sets and (7.1) can be rewritten as

$$\bigcup_{m=1}^{\infty} A_{m,n} = X. \tag{7.3}$$

The equation above implies that $\lim_{m\to\infty} \mu(A_{m,n}) = \mu(X)$. Thus there exists $m_n \in \mathbb{N}$ such that

$$\mu(X) - \mu(A_{m_n,n}) < \frac{\epsilon}{2^n}. \tag{7.4}$$

Now let

$$E = \bigcap_{n=1}^{\infty} A_{m_n,n}. \tag{7.5}$$

Then

$$\begin{aligned} \mu(X - E) &= \mu\left(X - \bigcap_{n=1}^{\infty} A_{m_n,n}\right) = \mu\left(\bigcup_{n=1}^{\infty} (X - A_{m_n,n})\right) \\ &\leq \sum_{n=1}^{\infty} \mu(X - A_{m_n,n}) < \epsilon. \end{aligned} \tag{7.6}$$

Now we show that $f_1, f_2, \ldots$ converges uniformly to f on E. To do this, suppose $\epsilon' > 0$. Let $n \in \mathbb{N}$ be such that $\frac{1}{n} < \epsilon'$. Then from $E \subset A_{m_n,n}$ we have

$$|f_k(x) - f(x)| < \frac{1}{n} < \epsilon'$$

for all $k \geq m_n$ and all $x \in E$. Hence $f_n \to f$ uniformly on E. □

Example Egorov's theorem does not hold, in general, in infinite measure spaces. For instance, consider the set $\mathbb{N}$ of natural numbers with the counting measure defined on the σ-algebra of all subsets of $\mathbb{N}$. If $F \subset \mathbb{N}$ is such that $\mu(F) < \epsilon < 1$, then, clearly, $F = \emptyset$. Thus uniform convergence on F^c means uniform convergence on $\mathbb{N}$. Now consider the sequence $\{f_n\}_{n=1}^{\infty}$ defined by $f_n = \chi_{\{1,2,\cdots,n\}}$. Then $f_n \to f$ on $\mathbb{N}$, where $f(i) = 1$ for all $i \in \mathbb{N}$, but this convergence is not uniform.

Definition 7.2 (simple function). Let X be a set. A function $s : X \to \mathbb{R}$ is called a simple function if it takes on only finitely many values, i.e., the image $s(X)$ is a finite subset of $\mathbb{R}$.

Suppose $(X, \mathcal{A}, \mu)$ is a measure space, $f : X \to \mathbb{R}$ is a simple function, and $c_1, \ldots, c_n$ are the distinct nonzero values of f. Then

$$f = c_1\chi_{E_1} + c_2\chi_{E_1} + \cdots + c_n\chi_{E_n},$$

where $E_i = f^{-1}(\{c_i\})$. Thus f is measurable if and only if $E_1, ..., E_n \in \mathcal{A}$.

Theorem 7.3 (approximation by simple functions) *Suppose $(X, \mathcal{A})$ is a measurable space and $f : X \to [-\infty, \infty]$ is measurable. Then there exists a sequence $f_1, f_2, ...$ of functions from X to $\mathbb{R}$ such that*

(a) each f_k is a simple measurable function;

(b) $|f_k(x)| \leq |f_{k+1}(x)| \leq |f(x)|$ for all $k \in \mathbb{N}$ and all $x \in X$;

(c) $\lim_{k\to\infty} f_k(x) = f(x)$ for every $x \in X$;

(d) $f_1, f_2, ...$ converges uniformly to f on any set on which f is bounded.

Proof. The idea of the proof is that for each $k \in \mathbb{N}$ and $m \in \mathbb{Z}$, the interval $[m, m+1)$ is divided into 2^k equally sized half-open subintervals. If $f(x) \in [0, k)$, we define $f_k(x)$ to be the left endpoint of the subinterval into which $f(x)$ falls; if $f(x) \in (-k, 0)$, we define $f_k(x)$ to be the right endpoint of the subinterval into which $f(x)$ falls; and if $|f(x)| \geq k$, we define $f_k(x)$ to be $\pm k$. Specifically, let

$$f_k(x) = \begin{cases} \frac{m}{2^k},\ if\ 0 \leq f(x) < k\ and\ m \in \mathbb{Z}\ is\ such\ that\ f(x) \in [\frac{m}{2^k}, \frac{m+1}{2^k}), \\ \frac{m+1}{2^k},\ if\ -k < f(x) < 0\ and\ m \in \mathbb{Z}\ is\ such\ that\ f(x) \in [\frac{m}{2^k}, \frac{m+1}{2^k}), \\ k,\ if\ f(x) \geq k, \\ -k,\ if\ f(x) \leq -k. \end{cases}$$

Each $f^{-1}([\frac{m}{2^k}, \frac{m+1}{2^k})) \in \mathcal{A}$ since f is a measurable function. Thus each f_k is an measurable simple function; in other words, (a) holds. Also, (b) holds because of how we have defined f_k . The definition of f_k implies that

$$|f_k(x) - f(x)| < \frac{1}{2^k}\ for\ all\ x \in X\ such\ that\ f(x) \in [-k, k]. \tag{7.7}$$

Thus we see that (c) holds.

Finally, (7.7) shows that (d) holds. □

Definition 7.4 Let $(X, \mathcal{A}, \mu)$ be a measure space and let $\{f_n\}_{n=1}^\infty$ be a sequence of real-valued measurable functions defined on X. Let f be a real-valued measurable function defined on X. We say that $\{f_n\}$ **converges in measure** to f, with notation $f_n \xrightarrow{\mu} f$, if for every $\epsilon > 0$, we have

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}) = 0. \tag{7.8}$$

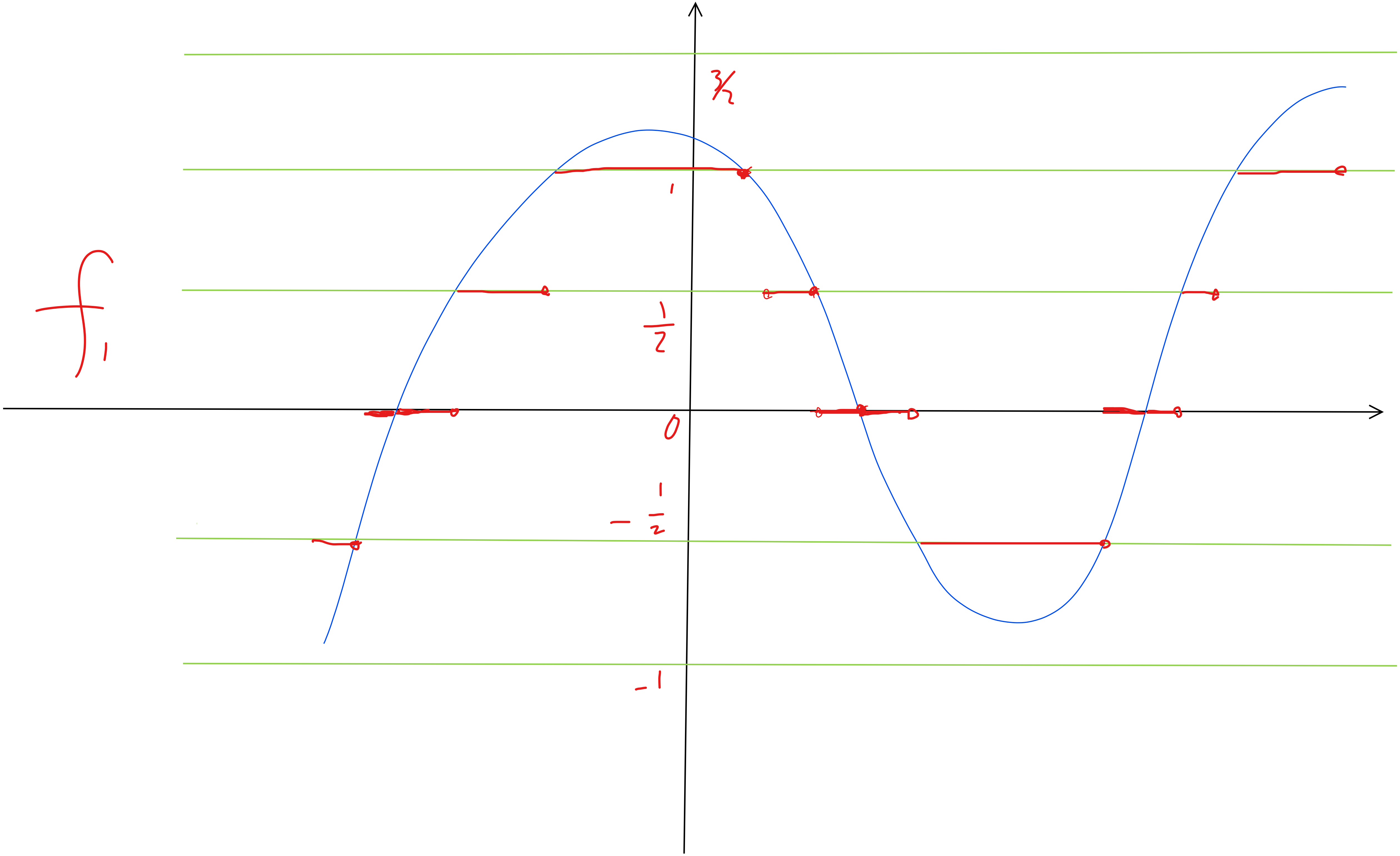
We say that $\{f_n\}$ is **Cauchy in measure** if for every $\epsilon > 0$ and for every $\delta > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, we have

$$\mu(\{x \in X : |f_n(x) - f_m(x)| \geq \epsilon\}) < \delta. \tag{7.9}$$

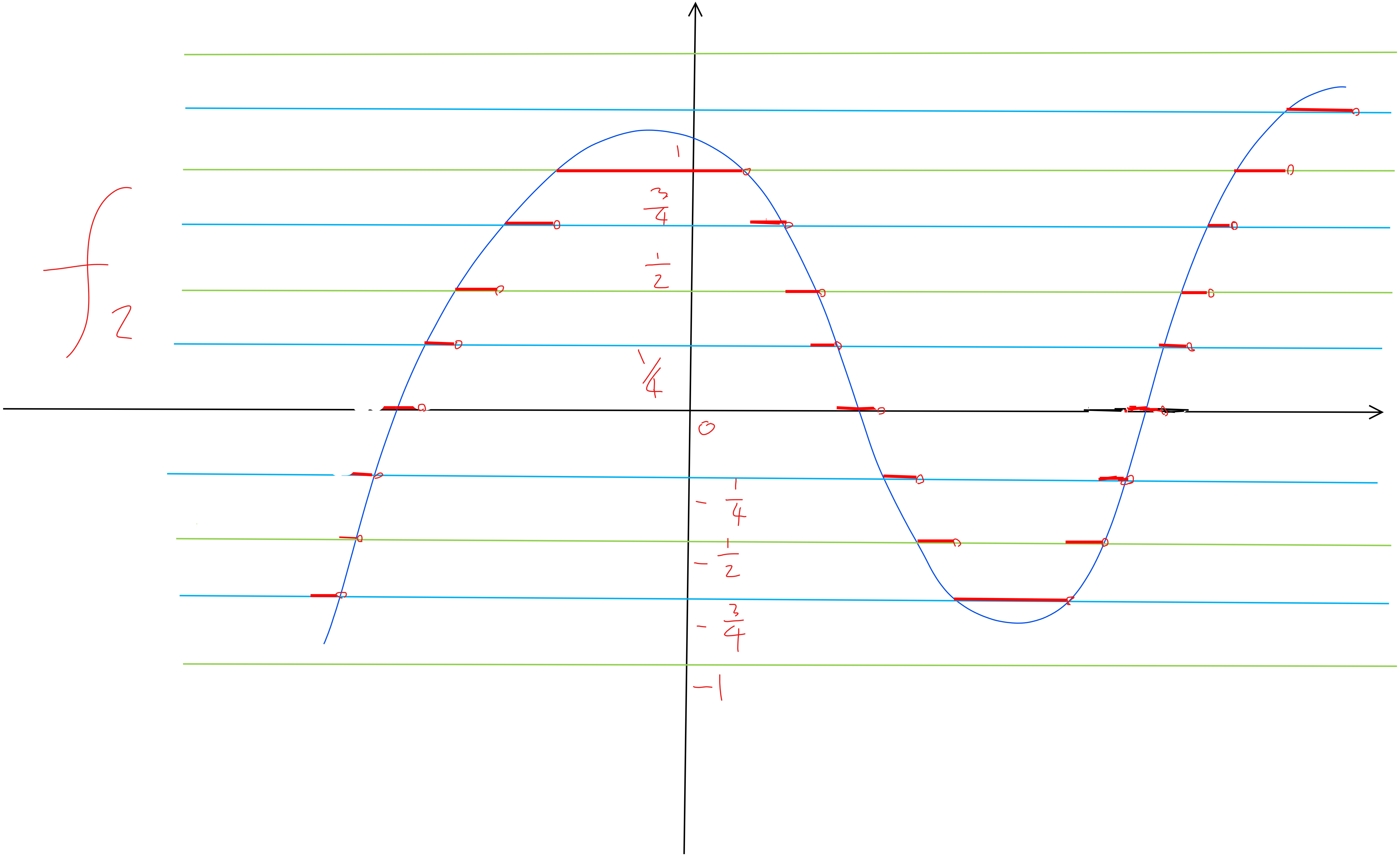
Proposition 7.5 *Let $(X, \mathcal{A}, \mu)$ be a measure space and $\mu(X) < \infty$. Let $\{f_n\}_{n=1}^\infty$ be a sequence of real-valued measurable functions, defined on X, converging a.e. to a real-valued measurable function f. Then $f_n \xrightarrow{\mu} f$.*

Proof. Let D denote the set of all points $x \in X$ such that the sequence $\{f_n(x)\}$ fails to converge to $f(x)$; then $\mu(D) = 0$. Let $\epsilon > 0$, and define sets $A_1, A_2, ...$ and $B_1, B_2, ...$ by

$$A_n = \{x \in X : |f_n(x) - f(x)| \geq \epsilon\}$$



f_1
$3/2$
1
$\frac{1}{2}$
0
$-\frac{1}{2}$
-1


f_2
1
$\frac{3}{4}$
$\frac{1}{2}$
$\frac{1}{4}$
0
$-\frac{1}{4}$
$-\frac{1}{2}$
$-\frac{3}{4}$
-1

and $B_n = \cup_{k=n}^{\infty} A_k$. The sequence $\{B_n\}$ is decreasing, and its intersection is included in D since

$$\begin{aligned} & \cap_{n=1}^{\infty} B_n = \limsup A_n \\ = & \ \{\text{those elemnts of X that belong to infinitely many } A_n's\}. \end{aligned}$$

Thus

$$0 = \mu(\cap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} \mu(B_n).$$

Since $A_n \subset B_n$, it follows that

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}) \leq \lim_{n\to\infty} \mu(B_n) = 0.$$

Hence $f_n \xrightarrow{\mu} f$. □

Example The above proposition is not true if $\mu(X) = \infty$. To see this, let $X = \mathbb{R}$ and let $f_n = \chi_{(n,n+1)}$. We have $f_n \to 0$ a.e., but f_n does not converge in measure to 0.

Proposition 7.6. *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let f and $f_1, f_2, ...$ be real-valued measurable functions on X. If $f_n \xrightarrow{\mu} f$, then there is a subsequence of $\{f_n\}$ that converges a.e. to f.*

Proof. The hypothesis that $\{f_n\}$ converges to f in measure means that

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}) = 0$$

holds for each $\epsilon > 0$. We use this relation to construct a sequence $\{n_k\}$ of positive integers, choosing n_1 so that

$$\mu(\{x \in X : |f_{n_1}(x) - f(x)| > 1\}) < \frac{1}{2},$$

and then choosing the remaining terms of $\{n_k\}$ inductively so that the relations $n_k > n_{k-1}$ and

$$\mu(\{x \in X : |f_{n_k}(x) - f(x)| > \frac{1}{k}\}) < \frac{1}{2^k}$$

hold for $k = 2, 3, ...$ Define sets $A_k, k = 1, 2, ...$, by

$$A_k = \{x \in X : |f_{n_k}(x) - f(x)| > \frac{1}{k}\}$$

If $x \notin \cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k$, then there is a positive integer j such that $x \notin \cup_{k=j}^{\infty} A_k$ and hence such that $|f_{n_k}(x) - f(x)| \leq \frac{1}{k}$ holds for $k = j, j+1, ...$ Thus $\{f_{n_k}\}$ converges to f at each x outside $\cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k$. Since

$$\mu(\cup_{k=j}^{\infty} A_k) \leq \sum_{k=j}^{\infty} \mu(A_k) \leq \sum_{k=j}^{\infty} \frac{1}{2^k} = \frac{1}{2^{j-1}}$$

holds for each j, it follows that $\mu(\cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k) = 0$, and the proof is complete. □

Exercises 6

1. Suppose f_n are measurable functions. Prove that

$$A = \{x : \lim_{n\to\infty} f_n(x) \textit{ exists } \}$$

is a measurable set.

2. Suppose $(X, \mathcal{A})$ is a measurable space, f is a real-valued function, and $\{x : f(x) > r\} \in \mathcal{A}$ for each rational number r. Prove that f is measurable.

3. Suppose $(X, \mathcal{A})$ is a measurable space. If $f : X \to \overline{\mathbb{R}}$ and $f^{-1}((r, \infty]) \in \mathcal{A}$ for each $r \in \mathbb{Q}$, then f is measurable.

4. Let $(X, \mathcal{A})$ be a measurable space. If $X = A \cup B$ where $A, B \in \mathcal{A}$. Show that a function f on X is measurable iff f is measurable on A and on B.

5. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{f_n\}, f$ measurable functions on X. Prove:

$$\{x \in X : \lim_{n\to\infty} f_n(x) \neq f(x)\} = \cup_{k=1}^{\infty} \cap_{m=k}^{\infty} \{x : \sup_{n>m} |f_n(x) - f(x)| > 1/k\}.$$

6. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $\{f_n\}$ finite a.e. measurable functions on X. Prove that there is a positive sequence $\{a_n\}$ such that $\lim_n a_n f_n(x) = 0$ a.e.

7. Let $(X, \mathcal{A}, \mu)$ be a measure space, $\{a_n\}$ a positive sequence, and $\{f_n\}$ measurable functions on X such that $\sum_n \mu(\{\frac{|f_n|}{a_n} > 1\}) < \infty$. Prove: a) $-1 \leq \liminf_n(|f_n|/a_n) \leq \limsup_n(|f_n|/a_n) \leq 1$ a.e. (b) If $a_n \to 0$, then $\lim_n f_n = 0$ a.e. and $\lim_n f_n = 0$ in measure.

8. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{f_n\}$ a monotone sequence of measurable functions on X that converges to f in measure. Prove that $\lim_{n\to\infty} f_n = f$ a.e.

9. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let f be a real-valued measurable function on a set $D \in \mathcal{A}$ with $\mu(D) < \infty$. Show that for every $\epsilon > 0$ there exists a set $E \in \mathcal{A}$ such that $E \subset D, \mu(D - E) < \epsilon$ and f is bounded on $D - E$.

10. Let f be an extended real-valued measurable function on a set $D \in \mathcal{A}$ in a measure space $(X, \mathcal{A}, \mu)$. For $M_1, M_2 \in \mathbb{R}, M_1 < M_2$, let the truncation off at M_1 and M_2 be defined by

$$g(x) = \begin{cases} M_1, \; if \; f(x) < M_1, \\ f(x), \; if \; f(x) \in [M_1, M_2], \\ M_2, \; if \; f(x) > M_2. \end{cases}$$

Show that g is measurable on D.

11. Let f be a bounded real-valued measurable function on a measure space $(X, \mathcal{A}, \mu)$.

(a) Show that there exists an increasing sequence of simple functions $\{f_n\}$, such that $f_n \uparrow f$ uniformly.

(b) Show that there exists a decreasing sequence of simple functions $\{g_n\}$, such that $g_n \downarrow f$ uniformly.

8. Integration

Integration is an important part in theory of measure. This section will introduce the integration on measure space and some convergence theorems which are the mail tools in the study of integration.

Definition 8.1 Let $(X, \mathcal{A}, \mu)$ be a measure space.
(i) Let

$$s = \sum_{i=1}^{n} a_i \chi_{E_i} \tag{8.1}$$

be a nonnegative measurable simple function. The (Lebesgue) integral of s over X is the number

$$\int_X s d\mu = \sum_{i=1}^{n} a_i \mu(E_i). \tag{8.2}$$

(ii) If $f \geq 0$ is a measurable function. The integral of f over X is the number $\int_X f du$ defined by

$$\int_X f d\mu := \sup_{s \leq f} \int_X s d\mu, \tag{8.3}$$

where the supremum is taken over all measurable simple functions $s : X \to \mathbb{R}$ satisfying $0 \leq s(x) \leq f(x)$ for all $x \in X$.

Let f be measurable and let $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. Provided $\int_X f^+ d\mu$ and $\int_X f^- d\mu$ are not both infinite, define

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu \tag{8.4}$$

Finally, if $f = u + iv$ is complex-valued and $\int_X (|u| + |v|) d\mu$ is finite, define

$$\int_X f d\mu = \int_X u d\mu + i \int_X v d\mu. \tag{8.5}$$

Remark A simple function s might be written in more than one way. For example $s = \chi_{A \cup B} = \chi_A + \chi_B$ if A and B are disjoint. It is not hard to check that the definition of $\int_X s d\mu$ is unaffected by how s is written. If

$$s = \sum_{i=1}^{n} a_i \chi_{A_i} = \sum_{j=1}^{m} b_j \chi_{B_j},$$

then we need to show

$$\sum_{i=1}^{n} a_i \mu(A_i) = \sum_{j=1}^{m} b_j \mu(B_j). \tag{8.6}$$

We leave the proof of this as an exercise.

The integral $\int_X f \chi_A d\mu$ is often written $\int_A f d\mu$.

Definition 8.2 If f is measurable and $\int_X |f| d\mu < \infty$, we say f is integrable.

Since $|f| = f^+ + f^-$, it now follows from the definition of integrability that both $\int_X f^+ d\mu$ and $\int_X f^- d\mu$ are finite.

The following proposition is an immediate consequence of the definition of the integral for non-negative functions.

Proposition 8.3 *Let $(X, \mathcal{A}, \mu)$ be a measure space and let f be a nonnegative, extended real-valued measurable function defined on X.*

(a) If g is a measurable function defined on X such that $0 \leq g \leq f$, and if $E \in \mathcal{A}$, then

$$\int_E g d\mu \leq \int_E f d\mu. \tag{8.7}$$

(b) If E and F are measurable subsets of X such that $E \subset F$, then

$$\int_E f d\mu \leq \int_F f d\mu. \tag{8.8}$$

(c) If c is a non-negative real number, and if E is a measurable subset of X, then

$$\int_E c f d\mu = c \int_E f d\mu. \tag{8.9}$$

(d) If E is a measurable subset of X such that $f(x) = 0$ for all $x \in E$, then

$$\int_E f d\mu = 0. \tag{8.10}$$

(e) If E is a measurable subset of X such that $\mu(E) = 0$, then

$$\int_E f d\mu = 0. \tag{8.11}$$

One of the most important results concerning integration is the monotone convergence theorem.

Theorem 8.4 *Suppose $\{f_n\}$ is a sequence of non-negative measurable functions with $f_1(x) \leq f_2(x) \leq \cdots$ for all x and with $\lim_{n\to\infty} f_n(x) = f(x)$ for all x. Then $\int_X f_n d\mu \to \int_X f d\mu$.*

Proof. By Proposition 8.3 (ii), $\{\int_X f_n d\mu\}$ is an increasing sequence. Let L be the limit. Since $f_n \leq f$ for all n, then $L \leq \int_X f d\mu$. We must show $L \geq \int_X f d\mu$.

Let $s = \sum_{i=1}^m a_i \chi_{E_i}$ be any non-negative simple function less than or equal to f and let $c \in (0, 1)$. Let $A_n = \{x : f_n(x) \geq cs(x)\}$. Since $\{f_n(x)\}$ increases to $f(x)$ for each x and $c < 1$, then $A_n \uparrow X$. For each n,

$$\begin{aligned} \int_X f_n d\mu &\geq \int_{A_n} f_n d\mu \geq c \int_{A_n} s d\mu \\ &= c \int_{A_n} \sum_{i=1}^m a_i \chi_{E_i} d\mu = c \sum_{i=1}^m a_i \mu(E_i \cap A_n). \end{aligned}$$

If we let $n \to \infty$, the right hand side converges to

$$c \sum_{i=1}^m a_i \mu(E_i) = c \int_X s d\mu.$$

Thus $L \geq c\int_X s d\mu$. Since c is arbitrary in the interval $(0,1)$, then $L \geq \int_X s d\mu$. Taking the supremum over all simple $s \leq f$, we obtain $L \geq \int_X f d\mu$. $\square$

Example 8.5 i) Let $X = [0,\infty)$ and $f_n(x) = -1/n$ for all x. Then $\int f_n = -\infty$, but $f_n \uparrow f$ where $f = 0$ and $\int f = 0$. The reason the monotone convergence theorem does not apply here is that the f_n are not non-negative.

ii) **Integration with respect to the counting measure** Let $\mathbb{N}$ be equipped with the counting measure. Let $f : \mathbb{N} \to \mathbb{R}$ be a given non-negative function. Let $f(k) = a_k \geq 0,\ k \in \mathbb{N}$. If we define

$$f_n(k) = \begin{cases} a_k, & \text{if } 1 \leq k \leq n, \\ 0, & \text{if } k > n, \end{cases}$$

then f_n increases to f. Notice that

$$f_n = \sum_{k=1}^{n} a_k \chi_{\{k\}}$$

is a non-negative simple function and so, by definition,

$$\int_{\mathbb{N}} f_n d\mu = \sum_{k=1}^{n} a_k \mu(\{k\}) = \sum_{k=1}^{n} a_k.$$

Thus, by the monotone convergence theorem, it follows that

$$\int_{\mathbb{N}} f d\mu = \sum_{k=1}^{\infty} a_k.$$

Once we have the monotone convergence theorem, we can prove that the Lebesgue integral is linear.

Theorem 8.6 *If f and g are non-negative and measurable or if f and g are integrable, then*

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu.$$

Proof. First suppose f and g are non-negative and simple, say,

$$f = \sum_{i=1}^{m} a_i \chi_{A_i},\ g = \sum_{j=1}^{n} b_j \chi_{B_j}$$

where $a_i, b_j \in [0,\infty)$ and $A_i, B_j \in \mathcal{A}$ are such that $A_i \cap A_{i'} = \emptyset$ for $i \neq i'$, $B_j \cap B_{j'} = \emptyset$ for $j \neq j'$, and $X = \cup_{i=1}^m A_i = \cup_{j=1}^n B_j$. Then

$$s + t = \sum_{i=1}^{m} \sum_{j=1}^{n} (a_i + b_j) \chi_{A_i \cap B_j},$$

and hence

$$\begin{aligned}\int_X (s+t)d\mu &= \sum_{i=1}^m \sum_{j=1}^n (a_i+b_j)\mu(A_i\cap B_j)\\ &= \sum_{i=1}^m a_i \sum_{j=1}^n \mu(A_i\cap B_j) + \sum_{j=1}^n b_j \sum_{i=1}^m \mu(A_i\cap B_j)\\ &= \sum_{i=1}^m a_i\mu(A_i) + \sum_{j=1}^n b_j\mu(B_j) = \int_X s d\mu + \int_X t d\mu.\end{aligned}$$

Thus the theorem holds in this case. Next suppose f and g are non-negative. Take s_n non-negative, simple, and increasing to f and t_n non-negative, simple, and increasing to g. Then s_n+t_n are simple functions increasing to $f+g$, so the result follows from the monotone convergence theorem and

$$\begin{aligned}\int_X (f+g)d\mu &= \lim_{n\to\infty}\int_X (s_n+t_n)d\mu\\ &= \lim_{n\to\infty}\int_X s_n d\mu + \lim_{n\to\infty}\int_X t_n d\mu = \int_X f d\mu + \int_X g d\mu.\end{aligned}$$

Suppose now that f and g are real-valued and integrable but take both positive and negative values. Since

$$\int |f+g| \le \int (|f|+|g|) = \int |f| + \int |g| < \infty,$$

then $f+g$ is integrable. Write

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

so that

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

Using the result for non-negative functions,

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

Rearranging,

$$\begin{aligned}\int (f+g) &= \int (f+g)^+ - \int (f+g)^-\\ &= \int f^+ - \int f^- + \int g^+ - \int g^- = \int f + \int g.\end{aligned}$$

If f and g are complex-valued, apply the above to the real and imaginary parts. □

The following result is left as an exercise.

Proposition 8.7 *Let $(X,\mathcal{A},\mu)$ be a measure space and let f and g be integrable, complex-valued functions defined on X. Let α and β be complex constants. Then*

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

Theorem 8.8 *(i) If s is a nonnegative simple function and $E_1, E_2, E_3, \ldots$ is a sequence of pairwise disjoint measurable sets, then*

$$\int_{\cup_{i=1}^{\infty} E_i} s d\mu = \sum_{i=1}^{\infty} \int_{E_i} s d\mu. \tag{8.12}$$

(ii) Let $f_n : X \to [0, \infty]$ be a sequence of measurable functions and define

$$f(x) = \sum_{i=1}^{\infty} f_i(x), \ \ x \in X.$$

Then $f : X \to [0, \infty]$ is measurable and for every $E \in \mathcal{A}$,

$$\int_E f d\mu = \sum_{i=1}^{\infty} \int_E f_i d\mu$$

Proof. (i) Let $E = \cup_{i=1}^{\infty} E_k, s = \sum_{k=1}^{m} a_k \chi_{A_k}$, then

$$\begin{aligned} \int_{\cup_{i=1}^{\infty} E_i} s d\mu &= \sum_{k=1}^{m} a_k \mu(A_k \cap E) = \sum_{k=1}^{m} a_k \sum_{i=1}^{\infty} \mu(A_k \cap E_i) \\ &= \sum_{k=1}^{m} a_k \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_k \cap E_i) \\ &= \lim_{n\to\infty} \sum_{k=1}^{m} a_k \sum_{i=1}^{n} \mu(A_k \cap E_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{k=1}^{m} a_k \mu(A_k \cap E_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \int_{E_i} s d\mu \\ &= \sum_{i=1}^{\infty} \int_{E_i} s d\mu. \end{aligned}$$

(ii) Let $F_n = \sum_{i=1}^{n} f_i$. For any $a \in \mathbb{R}$, since $F_n \uparrow f$, we have

$$\{x \in X : f(x) > a\} = \cup_{n=1}^{\infty} \{x \in X : F_n(x) > a\}.$$

Thus f is measurable and

$$\begin{aligned} \int_E f d\mu = \int_E \sum_{i=1}^{\infty} f_i d\mu &= \int_E \lim_{n\to\infty} \sum_{i=1}^{n} f_i d\mu = \int_E \lim_{n\to\infty} F_n d\mu \\ &= \lim_{n\to\infty} \int_E F_n d\mu = \sum_{i=1}^{\infty} \int_E f_n d\mu. \end{aligned}$$

□

Example Let $\{a_{ij}\}_{i,j=1}^{\infty}$ be a double sequence of non-negative real numbers and let $X = \mathbb{N}$ be equipped with the counting measure. Define

$$f_i(j) = a_{ij}, \ 1 \le i, j \le \infty.$$

Let $f = \sum_{i=1}^{\infty} f_i$; then

$$f(j) = \sum_{i=1}^{\infty} a_{ij}.$$

Now, by Theorem 8.8, we get

$$\int_X f d\mu = \sum_{i=1}^{\infty} \int_X f_i d\mu.$$

Using the result of Example 8.5 (ii), this translates into the following relation:

$$\sum_{j=1}^{\infty} f(j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_i(j).$$

Substituting the values for $f(j)$ and $f_i(j)$, we get

$$\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}.$$

Thus, for a non-negative double sequence of reals, the order of summation can be reversed. (Of course, both sums could be infinite.) This result is not true in general for sequences which change sign.

Theorem 8.9 (Fatou Lemma) *Suppose the f_n are non-negative and measurable. Then*

$$\int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n.$$

Proof. Let $g_n = \inf_{i\geq n} f_i$. Then the g_n are non-negative and g_n increases to $\liminf_{n\to\infty} f_n$. For each $i \geq n$, $g_n \leq f_i$ so $\int g_n \leq \int f_i$. Thus

$$\int g_n \leq \inf_{i\geq n} \int f_i.$$

If we take the limit in the above inequality as $n \to \infty$, on the left hand side we obtain $\int \liminf_{n\to\infty} f_n$ by the monotone convergence theorem, while on the right hand side we obtain $\liminf_{n\to\infty} \int_X f_n$. □

Example Let $(X, \mathcal{A}, \mu)$ be a measure space and $E \in \mathcal{A}$ be a measurable set such that $0 < \mu(E) < \mu(X)$. Defne $f_n = \chi_E$ when n is even and $f_n = 1 - \chi_E$ when n is odd. Then $\liminf_{n\to\infty} f_n = 0$ and so

$$\int_X \liminf_{n\to\infty} f_n = 0 < \min\{\mu(E), \mu(X - E)\} = \liminf_{n\to\infty} \int_X f_n.$$

Thus the inequality in Theorem 8.9 can be strict.

Theorem 8.10 (Dominated convergence theorem) *Suppose that f_n, f are measurable extended real-valued functions and $f_n(x) \to f(x)$ for each x. Suppose there exists a non-negative integrable function g such that $|f_n(x)| \leq g(x)$ for all x. Then*

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

Proof. Since $f_n + g \geq 0$, by Fatou's lemma,

$$\int f + \int g = \int (f+g) \leq \liminf_{n\to\infty} \int (f_n + g) = \liminf_{n\to\infty} \int f_n + \int g.$$

Since g is integrable,

$$\int f \leq \liminf_{n\to\infty} \int f_n. \tag{8.13}$$

Similarly, $g - f_n \geq 0$, so

$$\int g - \int f = \int (g-f) \leq \liminf_{n\to\infty} \int (g - f_n) = \int g + \liminf_{n\to\infty} \int (-f_n),$$

and so

$$-\int f \leq \liminf_{n\to\infty} \int (-f_n) = -\limsup_{n\to\infty} \int f_n, \tag{8.14}$$

which with (8.13) proves the theorem. □

Theorem 8.11 (An Extension of the Dominated Convergence Theorem) *Given a measure space $(X, \mathcal{A}, \mu)$. Let $\{f_n\}$ and f be extended real-valued measurable functions and let $\{g_n\}$ and g be nonnegative extended real-valued measurable functions on X. Suppose*

a) $f_n \to f$ and $g_n \to g$ a.e. on X.
b) g_n and g are all integrable and $\int_X g_n d\mu \to \int_X g d\mu$.
c) $|f_n| \leq g_n$ on X for any $n \in \mathbb{N}$.
Then f is integrable and $\int_X f_n d\mu \to \int_X f d\mu$.

Proof. Let us prove the integrability of f. We have

$$\begin{aligned}
\int_X |f| d\mu &= \int_X \lim_{n\to\infty} |f_n| d\mu \\
&\leq \liminf_{n\to\infty} \int_X |f_n| d\mu \\
&\leq \liminf_{n\to\infty} \int_X g_n d\mu \\
&= \lim_{n\to\infty} \int_X g_n d\mu \\
&= \int_X g d\mu < \infty.
\end{aligned}$$

This proves theintegrability of f on X.

Now $c)$ and $a)$ imply that $\{g_n + f_n\}$ and $\{g_n - f_n\}$ are sequences of nonnegative extended real-valued measurable functions and $\lim_{n\to\infty}(g_n + f_n) = g + f$ and $\lim_{n\to\infty}(g_n - f_n) = g - f$. Thus by Faton's Lemma we have

$$\int (g+f) \leq \liminf_{n\to\infty} \int (g_n + f_n) = \int g + \liminf_{n\to\infty} \int f_n, \tag{8.15}$$

and

$$\int (g-f) \leq \liminf_{n\to\infty} \int (g_n - f_n) = \int g + \liminf_{n\to\infty} \left(-\int f_n \right). \tag{8.16}$$

Now (8.15) implies
$$\int f \leq \liminf_{n\to\infty} \int f_n,$$
and (8.16) implies
$$\int f \geq \limsup_{n\to\infty} \int f_n$$
and consequently $\lim_{n\to\infty} \int f_n$ exists and $\lim_{n\to\infty} \int f_n = \int f$.

□

Theorem 8.12 *If $f : X \to [0, \infty]$ is measurable and $E_1, E_2, E_3, ...$ is a sequence of pairwise disjoint measurable sets. Then*
$$\int_{\cup_{i=1}^{\infty} E_i} f d\mu = \sum_{i=1}^{\infty} \int_{E_i} f d\mu. \tag{8.17}$$

Proof. Let
$$E = \cup_{i=1}^{\infty} E_i, \;\; f_n = \sum_{k=1}^{n} f\chi_{E_k};$$
then $f_n \uparrow f\chi_E$ and
$$\int_X f_n d\mu = \sum_{k=1}^{n} \int_{E_k} f d\mu. \tag{8.18}$$
Taking $n \to \infty$ and using the monotone convergence theorem, one gets (8.17).

Chebychev's Inequality *Let $(X, \mathcal{A}, \mu)$ be a measure space f and λ a positive real number.*

i) If f a nonnegative measurable function on X, then
$$\mu(\{x \in X, f(x) > \lambda\}) \leq \frac{1}{\lambda} \int_X f d\mu.$$
ii) If g is an integrable function on X, then
$$\mu(\{x \in X, |g(x)| > \lambda\}) \leq \frac{1}{\lambda} \int_{\{x \in X, |g(x)| > \lambda\}} |g| d\mu.$$

Proof. i) We have
$$\begin{aligned} \int_X f d\mu &= \int_{\{x\in X, f(x) > \lambda\}} f d\mu + \int_{\{x \in X, f(x) \leq \lambda\}} f d\mu \\ &\geq \int_{\{x\in X, f(x) > \lambda\}} f d\mu \geq \lambda \int_{\{x\in X, f(x) > \lambda\}} d\mu = \lambda\mu(\{x \in X, f(x) > \lambda\}). \end{aligned}$$

ii)
$$\int_{\{x\in X, |g(x)| > \lambda\}} |g| d\mu \geq \lambda \int_{\{x\in X, |g(x)| > \lambda\}} d\mu = \lambda\mu(\{x \in X, |g(x)| > \lambda\}).$$

□

Proposition 8.13 *Let $(X, \mathcal{A}, \mu)$ be a measure space and f a nonnegative measurable function on X for which $\int_X f d\mu < \infty$. Then f is finite a.e. on X and $\{x \in X : f(x) > 0\}$ is σ-finite.*

Proof. Let $X_\infty = \{x \in X : f(x) = \infty\}$; then

$$\int_X f d\mu = \int_{X_\infty} f d\mu + \int_{\{x \in X : f(x) < \infty\}} f d\mu \geq \int_{X_\infty} f d\mu \geq \int_{X_\infty} M d\mu = M\mu(X_\infty)$$

for any $M > 0$. Thus

$$\mu(X_\infty) \leq \frac{1}{M} \int_X f d\mu, \ \forall M > 0.$$

Taking $M \to \infty$, we obtain $\mu(X_\infty) = 0$. Hence, f is finite a. e.

Define $X_n = \{x \in X : f(x) > 1/n\}$. By Chebychev's Inequality,

$$\mu(X_n) \leq n \int_X f d\mu < \infty.$$

Since $\{x \in X : f(x) > 0\} = \cup_{n=1}^\infty X_n$, $\{x \in X : f(x) > 0\}$ is σ-finite.

Proposition 8.14 *Suppose f is measurable and non-negative and $\int_X f d\mu = 0$. Then $f = 0$ a.e.*

Proof. If f were not equal to 0 almost everywhere, there exists an n such that $\mu(A_n) > 0$ where $A_n = \{x : f(x) > 1/n\}$. But since f is non-negative,

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

a contradiction. $\square$

Proposition 8.15 *Suppose f is real-valued and integrable and for every measurable set A we have $\int_A f d\mu = 0$. Then $f = 0$ almost everywhere.*

Proof. Let $\epsilon > 0$, $A = \{x : f(x) > \epsilon\}$. Then

$$0 = \int_A f d\mu \geq \epsilon\mu(A).$$

Hence $\mu(A) = 0$. We use this argument for $\epsilon = 1/n$ and $n = 1, 2, ...$ to conclude

$$\mu(\{x : f(x) > 0\}) = \mu(\cup_{i=1}^\infty \{x : f(x) > 1/i\}) \leq \sum_{i=1}^\infty \mu(\{x : f(x) > 1/i\}) = 0.$$

Similarly, $\mu(\{x : f(x) < 0\}) = 0$.

Definition 8.16 Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{f_n\}$ a sequence of functions on X, each of which is integrable over X. The sequence $\{f_n\}$ is said to be *uniformly integrable* over X provided for each $\epsilon > 0$, there is a $\delta > 0$ such that for any natural number n and measurable subset E of X, if $\mu(E) < \delta$, then

$$\int_E |f_n| d\mu < \epsilon. \tag{8.19}$$

The sequence $\{f_n\}$ is said to be *tight* over X provided for each $\epsilon > 0$, there is a subset $X_0 \subset X$ that has finite measure and, for any natural number n,

$$\int_{X-X_0} |f_n| d\mu < \epsilon.$$

Proposition 8.17 *Let $(X, \mathcal{A}, \mu)$ be a measure space and the function f be integrable over X. Then for each $\epsilon > 0$, there is a $\delta > 0$ such that for any measurable subset E of X,*

$$\mu(E) < \delta \Rightarrow \int_E |f| d\mu < \epsilon. \tag{8.20}$$

Furthermore, for each $\epsilon > 0$, there is a subset $X_0 \subset X$ that has finite measure and

$$\int_{X-X_0} |f| d\mu < \epsilon. \tag{8.21}$$

Proof. We assume $f \geq 0$ on X. The general case follows by considering the positive and negative parts of f. Let $\epsilon > 0$. Since $\int_X f d\mu$ is finite, by the definition of the integral of a nonnegative function, there is a measurable simple function s on X for which

$$0 \leq s \leq f \text{ on } X \text{ and } 0 \leq \int_X f - \int_X s < \epsilon/2.$$

Choose $M > 0$ such that $0 \leq s \leq M$ on X. Therefore, by the linearity and monotonicity of integration, if $E \subset X$ is measurable, then

$$\int_E f = \int_E s + \int_E (f - s) \leq \int_E s + \frac{\epsilon}{2} \leq M\mu(E) + \frac{\epsilon}{2}.$$

Thus (8.20) holds for $\delta = \frac{\epsilon}{2M}$. Since the simple function s is integrable over X, the measurable set $X_0 = \{x \in X : s(x) > 0\}$ has finite measure. Moreover,

$$\int_{X-X_0} f = \int_{X-X_0} (f - s) \leq \int_X (f - s) < \epsilon. \tag{8.22}$$

The proof is complete. □

Vitali Convergence Theorem *Let $(X, \mathcal{A}, \mu)$ be a measure space and $\{f_n\}$ a sequence of functions on X that is both uniformly integrable and tight over X. Assume $f_n \to f$ pointwise a.e. on X and the function f is integrable over X. Then*

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

Proof. Observe that for all $n, |f - f_n| \leq |f| + |f_n|$ on X. If X_0 and X_1 are measurable subsets of X for which $X_1 \subset X_0$, then for all n, since X is the disjoint union $X = X_1 \cup [X_0 - X_1] \cup [X - X_0]$,

$$\left|\int_X (f_n - f)\right| \leq \int_{X_1} |f_n - f| + \int_{X_0-X_1} (|f_n| + |f|) + \int_{X-X_0} (|f_n| + |f|) \tag{8.23}$$

Let $\epsilon > 0$. By the preceding proposition, the tightness of $\{f_n\}$, and the linearity of integration, there is a measurable subset X_0 of X of finite measure for which

$$\int_{X-X_0} (|f_n| + |f|) = \int_{X-X_0} |f_n| + \int_{X-X_0} |f| < \frac{\epsilon}{3} \ for\ all\ n. \tag{8.24}$$

By the preceding proposition, the uniform integrability of $\{f_n\}$, and the linearity of integration, there is a $\delta > 0$ such that for any measurable subset E of X,

$$if\ \mu(E) < \delta\ then\ \int_E (|f_n| + |f|) = \int_E |f_n| + \int_E |f| < \frac{\epsilon}{3}.$$

By assumption, f is integrable over X. Thus f is finite a.e. on X. Moreover, from $\mu(X_0) < \infty$, we may apply Egorov's Theorem to infer that there is a measurable subset X_1 of X_0 for which $\mu(X_0 - X_1) < \delta$ and $\{f_n\}$ converge uniformly on X_1 to f. It follows from that

$$\int_{X_0-X_1} (|f_n| + |f|) < \frac{\epsilon}{3},\ for\ all\ n. \tag{8.25}$$

On the other hand, by the uniform convergence of $\{f_n\}$ to f on X_1, a set of finite measure, there is an N for which

$$\int_{X_1} |f_n - f| \le \sup_{x\in X_1} |f_n(x) - f(x)|\mu(X_1) < \frac{\epsilon}{3},\ for\ all\ n \ge N. \tag{8.26}$$

Combining (8.23)-(8.26), we conclude that

$$\left|\int_X (f_n - f)\right| < \epsilon\ for\ all\ n \ge N.$$

The proof is complete. □

The following result relates the Riemann integral and the Lebesgue integral and the proof of which is omitted.

Theorem 8.18 *A bounded real-valued function f on $[a, b]$ is Riemann integrable if and only if the set of points at which f is discontinuous has Lebesgue measure 0, and in that case, f is Lebesgue measurable and the Riemann integral of f is equal in value to the Lebesgue integral of f.*

Exercises 7

1. Suppose f is a non-negative integrable function on a measure space $(X, \mathcal{A}, \mu)$. Prove that $\lim_{t\to\infty} t\mu(\{x : f(x) \ge t\}) = 0$.

2. Let X be a set and $\mathcal{A} = P(X)$, the collection of all subsets of X. Pick $x \in X$ and let δ_x be the Dirac measure at x. Prove that if $f : X \to \mathbb{R}$, then

$$\int_X f d\delta_x = f(x).$$

3. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to [0, \infty]$ be a measurable function. Then the function

$$\mu_f : \mathcal{A} \to [0, \infty],$$

defined by

$$\mu_f(A) = \int_A f d\mu$$

is a measure, and

$$\int_E g d\mu_f = \int_E g f d\mu$$

for every measurable function $g : X \to [0, \infty]$ and every $E \in \mathcal{A}$.

4. If $\int_X f d\mu > 0$, where f is a measurable extended real-valued function on X, show that $\{x \in X : f(x) > 0\}$ has positive measure.

5. Let $\{f_n\}$ and $\{g_n\}$ be sequences of measurable functions on X such that $|f_n| \le g_n$ for every n. Let f and g be measurable functions such that $\lim_n f_n = f$ a.e. and $\lim_n g_n = g$ a.e. If $\lim_n \int_X g_n = \int_X g < \infty$, show that

$$\lim_n \int_X f_n = \int_X f.$$

6. Let $\{f_n\}$ be sequences of integrable functions on X such that $f_n \to f$ a.e., where f is also integrable. Show that $\lim_n \int_X |f_n - f| d\mu = 0$ iff $\lim_n \int_X |f_n| d\mu = \int_X |f| d\mu$.

7. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to [0, \infty]$ be measurable, $\int_X f d\mu = c \in (0, \infty)$, and a is a constant. Prove that

$$\lim_{n\to\infty} \int_X n \log(1 + (f/n)^a) d\mu = \begin{cases} \infty, \ if \ 0 < a < 1, \\ c, \ if \ a = 1 \\ 0, \ if \ 1 < a < \infty. \end{cases}$$

Hint: If $a \ge 1$, the integrands are dominated by af. If $a < 1$, Fatou's lemma can be applied.

8. Suppose $\mu(X) < \infty, \{f_n\}$ is a sequence of bounded complex measurable functions on X, and $f_n \to f$ uniformly on X. Prove that

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

9. Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that if $f \in L^1(\mu)$, then

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

10. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and f a measurable function on X. For an integer n, let $A_n = \{|f| > n\}$ and $B_n = \{n \le |f| < n + 1\}$. Prove that the following statements are equivalent: a) $f \in L^1(\mu)$; b) $\sum_n n\mu(B_n) < \infty$; $\sum_n \mu(A_n) < \infty$.

11. Let $(X, \mathcal{A}, \mu)$ be a measure space and f a nonnegative integrable function on X. Prove that for every $\epsilon > 0$, there is $A \in \mathcal{A}$ with $\mu(A) < \infty$ such that $\int_X f d\mu < \int_A f d\mu + \epsilon$.

12. Let $(X, \mathcal{A}, \mu)$ be a measure space, f an integrable function on X and $A \subset X$ with $\mu(A) < \infty$. Prove that for some $x \in A$,

$$|f(x)| \leq \frac{1}{\mu(A)} \int_A |f| d\mu.$$

13. Let $(X, \mathcal{A}, \mu)$ be a measure space and f a measurable function on X such that $\int_A f d\mu \geq 0$ for all $A \in \mathcal{A}$. Prove that $f \geq 0$ a.e. From this deduce that if $\int_A f d\mu = 0$ for all $A \in \mathcal{A}$, then $f = 0$ a.e.

14. Let $(X, \mathcal{A}, \mu)$ be a measure space and $\mathcal{F}$ be a π-system such that $\sigma(\mathcal{F}) = \mathcal{A}$. Suppose that f is integrable, or measurable and nonnegative μ-a.e. Prove that if $\int_A f d\mu = 0$ for all $A \in \mathcal{F}$ and $A = X$, then $f = 0$ μ-a.e.

15. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space, $F \subset \mathbb{R}$ a closed set, and f an integrable function on X such that

$$\frac{1}{\mu(A)} \int_A f d\mu \in F \textit{ for all } A \in \mathcal{A} \textit{ with } \mu(A) < \infty.$$

Prove that $f(x) \in F$ for a.e. $x \in X$.

16. Let $(X, \mathcal{A}, \mu)$ be a measure space and $f \in L^1(\mu)$ such that $|\int_A f d\mu| \leq c\mu(A)$ for all measurable A with $\mu(A) < \infty$ and a constant c. Prove that $|f| \leq c$ a.e.

17. Given a measure space $(X, \mathcal{A}, \mu)$. Let f be a bounded real-valued measurable function on a set $D \in \mathcal{A}$ with $\mu(D) < \infty$. Suppose $|f(x)| \leq M$ for $x \in D$ for some constant $M > 0$.

(a) Show that if $\int_D f d\mu = M\mu(D)$, then $f = M$ a.e.

(b) Show that if $f < M$ a.e. on D and if $\mu(D) > 0$, then $\int_D f d\mu < M\mu(D)$.

18. Give a proof of Proposition 8.7.

9. Signed measure, Hahn, Jordan decompositions, Radon-Nikodym theorem

Signed measures have the countable additivity property of measures, but are allowed to take negative as well as positive values.

Definition 9.1 Let $\mathcal{A}$ be a σ-algebra. A signed measure is a function $\mu : \mathcal{A} \to (-\infty, \infty]$ such that

i) $\mu(\emptyset) = 0$;

ii) if $A_1, A_2, ...$ are pairwise disjoint and all the A_i are in $\mathcal{A}$, then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

If the series in part ii) of Definition 9.1 were to be conditionally convergent, then it would be possible to rearrange it to converge to an arbitrary desired sum; however, the assertion of ii) implies that rearranging it does not alter the sum. Thus, it follows from ii) that the series cannot be conditionally convergent, which is to say, it must be either absolutely convergent or properly divergent to ∞.

Let μ be a signed measure on a measurable space $(X, \mathcal{A})$. As in the case of measure, if $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ and $A_j \uparrow \cup_{j=1}^{\infty} A_j$, then $\mu(\cup_{j=1}^{\infty} A_j) = \lim_{j\to\infty} \mu(A_j)$. Similarly, if $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ and $A_j \downarrow \cap_{j=1}^{\infty} A_j$, $\mu(A_1) < \infty$, then $\mu(\cap_{j=1}^{\infty} A_j) = \lim_{j\to\infty} \mu(A_j)$.

A signed measure μ on a measurable space $(X, \mathcal{A})$ is said to be finite if $\mu(X) < \infty$ and it is said to be σ-finite if there exists a sequence of sets $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ such that $X = \cup_{j=1}^{\infty} A_j$ and $\mu(A_j) < \infty$ for every j.

The measure defined before is sometimes called *a positive measure.*

Definition 9.2 Let μ be a signed measure. A set $A \in \mathcal{A}$ is called a positive set for μ if $\mu(B) \geq 0$ whenever $B \subset A$ and $B \in \mathcal{A}$. We say $A \in \mathcal{A}$ is a negative set if $\mu(B) \leq 0$ whenever $B \subset A$ and $B \in \mathcal{A}$. A *null set* A is one where $\mu(B) = 0$ whenever $B \subset A$ and $B \in \mathcal{A}$.

Note that if μ is a signed measure, then

$$\mu(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(\cup_{i=1}^{n} A_i).$$

The proof is the same as in the case of positive measures.

Example 9.3 Suppose m is the Lebesgue measure and let f be an integrable fuction. Let

$$\mu(A) = \int_A f dm.$$

If f takes both positive and negative values, then μ is a signed measure. If we let $P = \{x : f(x) \geq 0\}$, then P is easily seen to be a positive set, and if $N = \{x : f(x) < 0\}$, then N is a negative set. The Hahn decomposition which we shall prove soon is a decomposition of our space (in this case $\mathbb{R}$) into the positive and negative sets P and N. This decomposition is unique, except that $C = \{x : f(x) = 0\}$ could be included in N instead of P, or apportioned partially to P and partially to N.

Proposition 9.4 *Let μ be a signed measure which takes values in $(-\infty, \infty]$. Let E be measurable with $\mu(E) < 0$. Then there exists a measurable subset F of E that is a negative set with $\mu(F) < 0$.*

Proof. If E is a negative set, we are done. If not, there exists a measurable subset of E with positive measure. Let n_1 be the smallest positive integer such that there exists measurable $E_1 \subset E$ with $\mu(E_1) \geq 1/n_1$. We then define pairwise disjoint measurable sets $E_2, E_3, ...$ by induction as follows. Let $k \geq 2$ and suppose $E_1, ..., E_{k-1}$ are pairwise disjoint measurable subsets of E with $\mu(E_i) > 0$ for $i = 1, ..., k-1$. If $F_k = E - (E_1 \cup \cdots E_{k-1})$ is a negative set, from

$$\mu(F_k) = \mu(E) - \sum_{i=1}^{k-1} \mu(E_{k-1}) < 0,$$

we know that F_k is the desired set F. If F_k is not a negative set, let n_k be the smallest positive integer such that there exists $E_k \subset F_k$ with E_k measurable and $\mu(E_k) \geq 1/n_k$. This implies in particular that when $n_k > 1$ we have for any measurable $W \subset F_k$ that $\mu(W) < \frac{1}{n_k - 1}$.

We stop the construction if there exists k such that F_k is a negative set with $\mu(F_k) < 0$. If not, we continue and let $F = E - (\cup_{k=1}^{\infty} E_k)$. Since $0 > \mu(E) > -\infty$

and $\mu(E_k) > 0$, we have from

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

that $-\infty < \mu(F) \leq \mu(E) < 0$, so the sum converges. Thus $\mu(E_k) \to 0$, and so $n_k \to \infty$.

It remains to show that F is a negative set. Suppose $G \subset F$ is measurable, then $G \subset E - (E_1 \cup \cdots E_{k-1}) = F_k$ for any $k = 2, 3, \cdots$. Therefore, if k is large, then

$$\mu(G) < \frac{1}{n_k - 1}.$$

Taking $k \to \infty$, we get $\mu(G) \leq 0$. □

We write $A\Delta B$ for $(A - B) \cup (B - A)$. The following is known as the *Hahn decomposition*.

Theorem 9.5 *Let μ be a signed measure taking values in $(-\infty, \infty]$.*

(1) There exist disjoint measurable sets E and F in $\mathcal{A}$ whose union is X and such that E is a negative set and F is a positive set.

(2) If E_0 and F_0 are another such pair, then $E\Delta E_0 = F\Delta F_0$ is a null set with respect to μ.

(3) If μ is not a positive measure, then $\mu(E) < 0$. If $-\mu$ is not a positive measure, then $\mu(F) > 0$.

Proof. (1) If there is no subset of negative measure, there is nothing to prove. Suppose that there is a subset of negative measure, which can be assumed to be a negative set from Proposition 9.4. Let

$$L = \inf\{\mu(A) : A \text{ is a negative set}\}.$$

Then $L < 0$. Choose negative sets $\{A_n\}$ such that $\mu(A_n) \to L$. Let $E = \cup_{i=1}^{\infty} A_i$, $B_1 = A_1$ and $B_n = A_n - (B_1 \cup \cdots \cup B_{n-1})$ for each $n \geq 2$. Since A_n is a negative set and $B_n \subset A_n$, we know that B_n is also a negative set. Also, the B_n are disjoint and $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n = E$, $\cup_{i=1}^{n} B_i \uparrow \cup_{i=1}^{\infty} B_i$. If $C \subset E, C \in \mathcal{A}$, then

$$\mu(C) = \lim_{n\to\infty} \mu(C \cap (\cup_{i=1}^{n} B_i)) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(C \cap B_i) \leq 0. \tag{9.1}$$

Thus E is a negative set and so

$$L \leq \mu(E) = \mu(A_n) + \mu(E - A_n) \leq \mu(A_n).$$

Taking $n \to \infty$, we have $\mu(E) = L$. Hence $L > -\infty$.

Let $F = E^c$. If F were not a positive set, there would exist $B \subset F$ with $\mu(B) < 0$. By Proposition 9.4 there exists a negative set C contained in B with $\mu(C) < 0$. But then $E \cup C$ would be a negative set with $\mu(E \cup C) = \mu(E) + \mu(C) < \mu(E) = L$, a contradiction.

(2) To prove uniqueness, if E_0, F_0 are another such pair of sets and $A \subset E - E_0 \subset E$, then $\mu(A) \leq 0$. But $A \subset E - E_0 = F_0 - F \subset F_0$, so $\mu(A) \geq 0$. Therefore $\mu(A) = 0$. The same argument works if $A \subset E_0 - E$, and any subset of $E\Delta E_0$ can be written as the union of A_1 and A_2, where $A_1 \subset E - E_0$ and $A_2 \subset E_0 - E$.

(3) Suppose μ is not a positive measure but $\mu(E)=0$. Since

$$\mu(E)=\mu(A\cap E)+\mu(E-A)\le \mu(A\cap E),$$

if $A\in\mathcal{A}$, then

$$\mu(A)=\mu(A\cap E)+\mu(A\cap F)\ge \mu(E)+\mu(A\cap F)\ge 0.$$

which says that μ must be a positive measure, a contradiction. A similar argument applies for $-\mu$ and F. □

Let us say two measures μ and ν are *mutually singular* if there exist two disjoint sets E and F in $\mathcal{A}$ whose union is X with $\mu(E)=\nu(F)=0$. This is often written as $\mu\perp\nu$.

Example 9.6. If μ is Lebesgue measure restricted to [0,1/2], that is, $\mu(A)=m(A\cap[0,1/2])$, ν is Lebesgue measure restricted to $[1/2,1]$, $X=[0,1]$, $E=(1/2,1]$ and $F=[0,1/2]$ then μ and ν are mutually singular. This example works because the Lebesgue measure of $\{1/2\}$ is 0.

The following is known as the Jordan decomposition theorem.

Theorem 9.7 *If μ is a signed measure on a measurable space $(X,\mathcal{A})$, there exist positive measures μ^+ and μ^- such that $\mu=\mu^+-\mu^-$ and μ^+ and μ^- are mutually singular. This decomposition is unique.*

Proof Let E and F be negative and positive sets, resp., for μ so that $X=E\cup F$ and $E\cap F=\emptyset$. Let

$$\mu^+(A)=\mu(A\cap F),\ \mu^-(A)=-\mu(A\cap E).$$

This gives the desired decomposition.

If $\mu=\nu^+-\nu^-$ is another such decomposition with ν^+, ν^- mutually singular, let E_1 and F_1 be sets such that $\nu^+(E_1)=0=\nu^-(F_1)$, $X=E_1\cup F_1$ and $E_1\cap F_1=\emptyset$. If $A\subset F_1$, then $\nu^-(A)\le\nu^-(F_1)=0$, and so

$$\mu(A)=\nu^+(A)-\nu^-(A)=\nu^+(A)\ge 0,$$

and consequently F_1 is a positive set for μ. Similarly, E_1 is a negative set for μ. Thus E_1,F_1 gives another Hahn decomposition of X. By the uniqueness part of the Hahn decomposition theorem, $F\Delta F_1$ is a null set with respect to μ. Since $\nu^+(E_1)=0$ and $\nu^-(F_1)=0$, if $A\in\mathcal{A}$, then

$$\begin{aligned}\nu^+(A) &= \nu^+(A\cap F_1)=\nu^+(A\cap F_1)-\nu^-(A\cap F_1)\\ &= \mu(A\cap F_1)=\mu(A\cap(F_1\cap F))+\mu(A\cap(F_1-F))\\ &= \mu(A\cap(F_1\cap F))+\mu(A\cap(F-F_1))=\mu(A\cap F)=\mu^+(A),\end{aligned}$$

and similarly $\nu^-=\mu^-$. □

Rmark. Since μ takes values in $(-\infty,\infty]$, from the construction of $\mu^\pm$ in the proof of Theorem 9.7, we know that μ^- is a finite measure. Thus if μ is a finite (respectively, σ-finite), then the same is true for $\mu^\pm$.

The measure

$$|\mu|=\mu^++\mu^-$$

is called the total variation measure of μ and $|\mu|(X)$ is called the total variation of X.

Suppose f is non-negative and integrable with respect to a measure μ. If we define ν by

$$\nu(A) = \int_A f d\mu, \tag{9.2}$$

then ν is a measure. The only part that needs thought is the countable additivity. If A_n are disjoint measurable sets, we have

$$\nu(\cup_{n=1}^{\infty} A_n) = \int_{\cup_{n=1}^{\infty} A_n} f d\mu = \sum_{n=1}^{\infty} \int_{A_n} f d\mu = \sum_{n=1}^{\infty} \nu(A_n)$$

by using (8.10). Moreover, $\nu(A)$ is zero whenever $\mu(A)$ is.

We consider the converse now. If we are given two measures μ and ν, when does there exist f such that (9.2) holds ? The Radon-Nikodym theorem answers this question.

Definition 9.9 A measure ν is said to be absolutely continuous with respect to a measure μ if $\mu(A) = 0 \Rightarrow \nu(A) = 0$. In this case, we write $\nu \ll \mu$.

Examples

(a) Let $(X, \mathcal{A}, \mu)$ be a measure space and f be an extended nonnegative real-valued measurable function on X. Define

$$\nu(E) = \int_E f d\mu, \ E \in \mathcal{A}.$$

Then as we have seen above, ν is a measure on $(X, \mathcal{A})$. Moreover, for all $E \in \mathcal{A}, \mu(E) = 0$ implies $\nu(E) = 0$. Consequently, $\nu \ll \mu$.

(b) Let $(\mathbb{R}, \mathcal{L}, m)$ be the Lebesgue measure space and ν be the counting measure on $\mathcal{L}$. If $\nu(E) = 0$, then $E = \emptyset$ and hence $m(E) = 0$. This shows that $m \ll \nu$.

(c) Let $(\mathbb{R}, \mathcal{L}, m)$ be the Lebesgue measure space and $\nu : \mathcal{L} \to [0, \infty]$ be defined by $\nu(\emptyset) = 0, \nu(E) = \infty$ for $E \in \mathcal{L}$ and $E \neq \emptyset$. Clearly, $m \ll \nu$.

(d) Let $X = \mathbb{N}$ and $\mathcal{A} = \mathcal{P}(\mathbb{N})$. For $E \in \mathcal{A}$, define $\mu(E) = \sum_{n \in E} 2^n$ and $\nu(E) = \sum_{n \in E} 2^{-n}$. Obviously, μ and ν are measures on $\mathcal{P}(\mathbb{N})$. Observe that $\mu(E) = 0$ if and only if $\nu(E) = 0$ if and only if $E = \emptyset$. This shows that $\mu \ll \nu$. and also $\nu \ll \mu$.

Proposition 9.10 *Let ν be a finite measure. Then ν is absolutely continuous with respect to μ if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that $\mu(A) < \delta$ implies $\nu(A) < \epsilon$.*

Proof. Suppose for each $\epsilon > 0$, there exists $\delta > 0$ such that $\mu(A) < \delta$ implies $\nu(A) < \epsilon$. If $\mu(A) = 0$, then $\nu(A) < \epsilon$ for all $\epsilon > 0$, hence $\nu(A) = 0$, and thus $\nu \ll \mu$. Suppose now that $\nu \ll \mu$. If there exists an $\epsilon > 0$ for which no corresponding δ exists, then there exists E_k such that $\mu(E_k) < \frac{1}{2^k}$ but $\nu(E_k) \geq \epsilon$. Let $F = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k$. Since

$$\cup_{k=n}^{\infty} E_k \downarrow F, \ \mu(\cup_{k=1}^{\infty} E_k) \leq \sum_{k=1}^{\infty} \frac{1}{2^k} < \infty,$$

one has

$$\mu(F) = \lim_{n\to\infty} \mu(\cup_{k=n}^{\infty} E_k) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(E_k) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} \frac{1}{2^k} = 0.$$

But

$$\nu(F) = \lim_{n\to\infty} \nu(\cup_{k=n}^{\infty} E_k) \ge \epsilon;$$

ν being finite is needed for the equality in the last line. This contradicts the absolute continuity. □

Lemma 9.11 *Let μ and ν be finite positive measures on a measurable space $(X, \mathcal{A})$. Then either $\mu \perp \nu$ or else there exist $\epsilon > 0$ and $G \in \mathcal{A}$ such that $\mu(G) > 0$ and G is a positive set for $\nu - \epsilon\mu$.*

Proof. Consider the Hahn decomposition for $\nu - \frac{1}{n}\mu$. Thus there exist a negative set E_n and a positive set F_n for this measure and $E_n \cap F_n = \emptyset$, $X = E_n \cup F_n$. Let $F = \cup_{n=1}^{\infty} F_n$ and $E = \cap_{n=1}^{\infty} E_n$. Note that

$$E^c = \cup_{n=1}^{\infty} E_n^c = \cup_{n=1}^{\infty} F_n = F.$$

For each $n, E \subset E_n$, so

$$\nu(E) \le \nu(E_n) \le \frac{1}{n}\mu(E_n) \le \frac{1}{n}\mu(X).$$

Since ν is a positive measure, this implies $\nu(E) = 0$.

One possibility is that $\mu(E^c) = 0$, in which case $\mu \perp \nu$. The other possibility is that $\mu(E^c) > 0$. In this case, $\mu(F_n) > 0$ for some n. Let $\epsilon = 1/n$ and $G = F_n$. Then from the definition of F_n, $\mu(G) > 0$ and G is a positive set for $\nu - \epsilon\mu$. □

The following is the *Radon-Nikodym theorem*.

Theorem 9.12 *Suppose μ and ν are σ-finite positive measures on a measurable space $(X, \mathcal{A})$ such that ν is absolutely continuous with respect to μ. Then there exists a μ-measurable non-negative function f such that*

$$\nu(A) = \int_A f d\mu \tag{9.3}$$

for all $A \in \mathcal{A}$. Moreover, if g is another such function, then $f = g$ almost everywhere with respect to μ.

The function f is called the Radon-Nikodym derivative of ν with respect to μ or sometimes the density of ν with respect to μ, and is written $f = d\nu/d\mu$. Sometimes one writes $d\nu = f d\mu$. The idea of the proof is to look at the set of f such that $\int_A f d\mu \le \nu(A)$ for each $A \in \mathcal{A}$, and then to choose the one such that $\int_X f d\mu$ is largest.

Proof. **Step 1**. Let us first prove the uniqueness assertion. Suppose f and g are two functions such that

$$\int_A f d\mu = \nu(A) = \int_A g d\mu$$

for all $A \in \mathcal{A}$. For every set A we have

$$\int_A (f-g)d\mu = \nu(A) - \nu(A) = 0.$$

Thus we have $f - g = 0$ a.e. with respect to μ.

Step 2. Let us assume μ and ν are finite measures for now. In this step we define the function f. Let

$$\mathcal{F} = \left\{ g \text{ measurable}: \ g \geq 0, \ \int_A g d\mu \leq \nu(A) \text{ for all } A \in \mathcal{A} \right\}.$$

$\mathcal{F}$ is not empty because $0 \in \mathcal{F}$. Let $L = \sup\{\int_X g d\mu : \ g \in \mathcal{F}\}$, note $L \leq \mu(X) < \infty$, and let g_n be a sequence in $\mathcal{F}$ such that $\int_X g_n d\mu \to L$. Let $h_n = \max\{g_1, ..., g_n\}$.

We *claim* that if $g_1, g_2 \in \mathcal{F}$, then $h_2 = \max\{g_1, g_2\} \in \mathcal{F}$. To see this, let $B = \{x : \ g_1(x) \geq g_2(x)\}$, and write

$$\begin{aligned} \int_A h_2 d\mu &= \int_{A\cap B} h_2 d\mu + \int_{A\cap B^c} h_2 d\mu \\ &= \int_{A\cap B} g_1 d\mu + \int_{A\cap B^c} g_2 d\mu \\ &\leq \nu(A\cap B) + \nu(A \cap B^c) = \nu(A). \end{aligned} \tag{9.4}$$

Therefore $h_2 \in \mathcal{F}$. By an induction argument, h_n is in $\mathcal{F}$.

The h_n increase, say to f. By the monotone convergence theorem,

$$\int_A f d\mu \leq \nu(A)$$

for all $A \in \mathcal{A}$ and

$$\int_X f d\mu \geq \int_X h_n d\mu \geq \int_X g_n d\mu$$

for each n, so $\int_X f d\mu = L$.

Step 3. Next we prove that f is the desired function. Define a measure λ by

$$\lambda(A) = \nu(A) - \int_A f d\mu.$$

λ is a positive measure since $f \in \mathcal{F}$. Suppose λ is not mutually singular to μ. By Lemma 9.1, there exists $\epsilon > 0$ and $G \in \mathcal{A}$ such that $\mu(G) > 0$, and G is a positive set for $\lambda - \epsilon\mu$. For any $A \in \mathcal{A}$,

$$\nu(A) - \int_A f d\mu = \lambda(A) \geq \lambda(A \cap G) \geq \epsilon\mu(A\cap G) = \int_A \epsilon\chi_G d\mu,$$

or

$$\nu(A) \geq \int_A (f + \epsilon\chi_G) d\mu.$$

Hence $f + \epsilon\chi_G \in \mathcal{F}$. But

$$\int_X (f + \epsilon\chi_G) d\mu = L + \epsilon\mu(G) > L,$$

a contradiction to the definition of L. Therefore $\lambda \perp \mu$. Then there must exist $H \in \mathcal{A}$ such that $\mu(H) = 0$ and $\lambda(H^c) = 0$. Since $\nu \ll \mu$, then $\nu(H) = 0$, and hence

$$\lambda(H) = \nu(H) - \int_H f d\mu = 0.$$

This implies $\lambda = 0$, or $\nu(A) = \int_A f d\mu$ for all A.

Step 4. General case. Since μ and ν are σ-finite measures, there exist sequences of measurable sets $A_n, B_n \in \mathcal{A}$ such that $A_n \uparrow X, B_n \uparrow X$, and $\mu(A_n) < \infty, \nu(B_n) < \infty$ for all n. Define $X_n := A_n \cap B_n$. Then $X_n \uparrow X$ and $\mu(X_n) < \infty, \nu(X_n) < \infty$ for all n. From the above proof, there exist a sequence of measurable functions $f_n : X_n \to [0, \infty)$ such that

$$\nu(A) = \int_A f_n d\mu \ for\ all\ n \in \mathbb{N}\ and\ all\ A \in \mathcal{A}\ such\ that\ A \subset X_n. \tag{9.5}$$

It follows from **Step 1** that the restriction of f_{n+1} to X_n agrees with f_n μ-almost everywhere. Thus, modifying f_{n+1} on a set of measure zero if necessary, we may assume without loss of generality that $f_{n+1}|_{X_n} = f_n$ for all $n \in \mathbb{N}$. With this understood, define

$$f : X \to [0, \infty)$$

by

$$f|_{X_n} := f_n \quad for\ n \in \mathbb{N}.$$

This function is measurable because

$$f^{-1}([0, c]) = \cup_{n=1}^{\infty}(X_n \cap f^{-1}([0, c])) = \cup_{n=1}^{\infty} f_n^{-1}([0, c]) \in \mathcal{A}.$$

Let $E \in \mathcal{A}$ and define $E_n := E \cap X_n \in \mathcal{A}, n \in \mathbb{N}$. Then

$$E_n \uparrow E.$$

It follows that

$$\begin{aligned} \nu(E) &= \lim_{n\to\infty} \nu(E_n) = \lim_{n\to\infty} \int_{E_n} f d\mu \\ &= \lim_{n\to\infty} \int_X f\chi_{E_n} d\mu \overset{MCT}{=} \int_X f\chi_E d\mu = \int_E f d\mu. \end{aligned}$$

□

Example 9.13. Let X be a one-element set and let $\mathcal{A} = \mathcal{P}(X)$. Define the measure $\mu : \mathcal{A} \to [0, \infty]$ by $\mu(\emptyset) = 0, \mu(X) = \infty$.

Choose $\nu(\emptyset) = 0$ and $\nu(X) = 1$. Then $\nu \ll \mu$, but there does not exist a (measurable) function $f : X :\to [0, \infty]$ such that $\int_X f d\mu = \nu(X)$. Thus the hypothesis that $(X, \mathcal{A}, \mu)$ is σ-finite cannot be removed in Theorem 9.12.

Define integration with respect to a signed measure μ by

$$\int_X f d\mu = \int_X f d\mu^+ - \int_X f d\mu^-,$$

where μ^+ and μ^- form the Jordan decomposition of μ.

Let f be an extended real-valued measurable function on a measure space $(X, \mathcal{A}, \mu)$ such that one of $\int_X f^+ d\mu$ and $\int_X f^- d\mu$ is finite, so that $\int_X f d\mu$ is defined as an extended real number. Set

$$\nu(E) = \int_E f d\mu, \text{ for } E \in \mathcal{A}.$$

Then ν is a signed measure on $\mathcal{A}$.

It is desired to determine when a signed measure is expressible as an integral, as is the case with ν in (9.3). The essential concept is absolute continuity and the main result is the Radon–Nikodm Theorem.

Definition 9.14 Let $(X, \mathcal{A})$ be a measurable space and μ and ν be signed measures on X. We say that ν is absolutely continuous with respect to μ if $|\mu|(A) = 0 \Rightarrow \nu(A) = 0$. In this case, we also write $\nu \ll \mu$.

The following result will be used later.

Theorem 9.15 (Radon-Nikodym) *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and ν be a σ-finite signed measure defined on $\mathcal{A}$ such that $\nu \ll \mu$. Then, there exists a measurable function f defined on X such that*

$$\nu(E) = \int_E f d\mu$$

for every $E \in \mathcal{A}$. The function f is unique in the sense that if g is any real-valued measurable function on X with $\nu(E) = \int_E g d\mu$ for all $E \in \mathcal{A}$, then $f = g$ μ-a.e.

Proof. Let $\nu = \nu^+ - \nu^-$ be the Jordan decomposition of ν. That is, $X = E \cup F$, $E \cap F = \emptyset$, E is a negative set, F is a positive set, $\nu^+(A) = \nu(A \cap F)$ and $\nu^-(A) = -\nu(A \cap E)$ for $A \in \mathcal{A}$. From the σ-finiteness of ν, we can write $X = \cup_{i=1}^{\infty} X_i$, where $X_i \in \mathcal{A}$ and $\nu(X_i) < \infty$. Observe that $\nu^-(X_j) = -\nu(X_j \cap E) < \infty$. Also, it follows from $\nu(X_j) = \nu(X_j \cap F) + \nu(X_j \cap E)$ and $\nu(X_j \cap E) > -\infty$ that $\infty > \nu(X_j \cap F) = \nu^+(X_j)$. Therefore, both ν^+ and ν^- are σ-finite. Since $\nu \ll \mu$, the positive measures ν^+ and ν^- satisfy $\nu^+ \ll \mu$ and $\nu^- \ll \mu$. From theorem 9.12, there exist nonnegative real-valued measurable f_1 and f_2 such that

$$\nu^+(E) = \int_E f_1 d\mu, \quad \nu^-(E) = \int_E f_2 d\mu$$

for every $E \in \mathcal{A}$. Since f_1 and f_2 are real-valued, their difference $f = f_1 - f_2$ is also real-valued. We **claim** that the inequalities $f^+ \leq f_1$ and $f^- \leq f_2$ hold. Consider any $x \in X$. If $f(x) \geq 0$ then $f_1(x) = f(x) + f_2(x) \geq f(x) = f^+(x)$, while if $f(x) \leq 0$ then $f_1(x) \geq 0 = f^+(x)$. Hence $f^+ \leq f_1$ on X. We deduce from here that $f_2(x) = f_1(x) - f(x) \geq f^+(x) - f(x) = f^-(x)$ for all $x \in X$. Thus our **claim** is true. Since $\nu^-(X) = -\nu(E) < \infty$, we have $\int_X f_2 d\mu < \infty$. In conjunction with the inequalities proved in the preceding paragraph, this implies that the function $f = f_1 - f_2$ has the property that $\int_X f^- d\mu < \infty$, and hence that $\int_X f d\mu$ is well defined (though not necessarily finite). Consequently, the equality $\int_E f d\mu = \int_E f_1 d\mu - \int_E f_2 d\mu$ holds for every $E \in \mathcal{A}$, which leads to

$$\nu(E) = \nu^+(E) - \nu^-(E) = \int_E f_1 d\mu - \int_E f_2 d\mu = \int_E f d\mu.$$

□

Exercises 8

1. Suppose μ is a signed measure. Prove that A is a null set with respect to μ if and only if $|\mu|(A) = 0$.

2. Let μ be a signed measure on a measurable space $(X, \mathcal{A})$ and f be a measurable function on X. Prove that

$$\left|\int_E f d\mu\right| \leq \int_E |f| d|\mu|, \ \forall \ E \in \mathcal{A}.$$

3. Let μ be a signed measure on $(X, \mathcal{A})$. Prove that if $A \in A$, then

$$|\mu|(A) = \sup\left\{\sum_{j=1}^{n} |\mu(B_j)|, each\ B_j \in \mathcal{A},\ the\ B_j\ are\ disjoint,\ \cup_{j=1}^{n} B_j = A\right\}.$$

4. If $\mu = \mu_1 - \mu_2$, where μ_1, μ_2 are positive measures and either μ_1 or μ_2 is finite, then $\mu_1(E) \geq \mu^+(E)$, $\mu_2(E) \geq \mu^-(E)$, for all $E \in \mathcal{A}$, where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ.

5. Let $(X, \mathcal{A}, \mu)$ be a finite signed measure space. Show that there is an $M > 0$ such that $|\mu(E)| < M$ for every $E \in \mathcal{A}$.

6. Show that a signed measure μ on $(X, \mathcal{A})$ is finite [resp. σ-finite] if and only if $|\mu|$ is finite [resp. σ-finite] if and only if $|\mu^+|$ and $|\mu^-|$ are both finite [resp. σ-finite].

7. Let $(X, \mathcal{A}, \mu)$ be a signed measure space. Then, for every $E \in \mathcal{A}$; the following hold:
(a) $\mu^+(E) = \sup\{\mu(F) : F \subset E, F \in \mathcal{A}\}$;
(b) $\mu^-(E) = \sup\{-\mu(F) : F \subset E, F \in \mathcal{A}\}$.

8. Let μ_1 and μ_2 be positive measures on $(X, \mathcal{A})$ such that for all $a > 0$ and $b > 0$, there exist sets $A_{a,b}, B_{a,b} \in \mathcal{A}$ satisfying $A_{a,b} \cup B_{a,b} = X, \mu_1(A_{a,b}) < a$, $\mu_2(B_{a,b}) < b$. Show that $\mu_1 \perp \mu_2$.

9. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and let ν be a measure on $(X, \mathcal{A})$ for which the conclusion of the Radon–Nikodym Theorem holds. Prove that ν is σ-finite.

10. Let μ and ν be measures on the measurable space $(X, \mathcal{A})$. Show that the absolute continuity condition is equivalent to $\mu(A \Delta B) = 0 \Rightarrow \nu(A) = \nu(B)$ for all $A, B \in \mathcal{A}$.

11. Let μ and ν be signed measures on a measurable space $(X, \mathcal{A})$. Show that the following conditions are equivalent:

$$(a)\ \nu \ll \mu\ (b)\ \nu^+ \ll \mu, \nu^- \ll \mu\ (c)\ |\nu| \ll |\mu|\ (d)\ |\nu| \ll \mu.$$

10. L^p spaces

Integration is a basic tool in the study of real analysis. This chapter considers the most important spaces in integration, the L^p spaces associated to a measure space $(X, \mathcal{A}, \mu)$.

Let $(X, \mathcal{A}, \mu)$ be a measure space and that p and q are real numbers such that

$$\frac{1}{p} + \frac{1}{q} = 1,\ 1 < p < \infty,\ 1 < q < \infty. \tag{10.1}$$

Any two nonnegative real numbers a and b satisfy Young's inequality

$$ab \le \frac{1}{p} a^p + \frac{1}{q} b^q \tag{10.2}$$

and equality holds in if and only if $a^p = b^q$. This can be proven by examining the critical points of the function

$$(0, \infty) \to \mathbb{R}:\ x \to \frac{1}{p} x^p - xb.$$

Theorem 10.1 *Let $f, g : X \to [0, \infty]$ be measurable functions. Then f and g satisfy the Hölder inequality*

$$\int_X fg d\mu \le \left(\int_X f^p d\mu \right)^{1/p} \left(\int_X g^q d\mu \right)^{1/q} \tag{10.3}$$

and the Minkowski inequality

$$\left(\int_X (f+g)^p d\mu \right)^{1/p} \le \left(\int_X f^p d\mu \right)^{1/p} + \left(\int_X g^q d\mu \right)^{1/q}. \tag{10.4}$$

Proof. (i) Define

$$A = \left(\int_X f^p d\mu \right)^{1/p},\ B = \left(\int_X g^q d\mu \right)^{1/q}.$$

We can assume $0 < A < \infty$ and $0 < B < \infty$ otherwise it is trivial. We can also assume that $A = B = 1$. Then it follows from (10.2) that

$$\int_X fg d\mu \le \int_X \left(\frac{f^p}{p} + \frac{g^q}{q} \right) d\mu = 1 = \left(\int_X f^p d\mu \right)^{1/p} \left(\int_X g^q d\mu \right)^{1/q}. \tag{10.5}$$

This proves the Hölder inequality. To prove the Minkowski inequality, put

$$C = \left(\int_X (f+g)^p d\mu \right)^{1/p}.$$

We must prove that $C \le A + B$. This is obvious when a $A = \infty$ or $B = \infty$. Hence assume $A, B < \infty$. We first show that $C < \infty$. This holds because

$$f \le (f^p + g^p)^{1/p},\ g \le (f^p + g^p)^{1/p},$$

hence

$$f + g \le 2(f^p + g^p)^{1/p},$$

thus

$$(f+g)^p \le 2^p (f^p + g^p)$$

and integrating this inequality and raising the integral to the power $1/p$ we obtain

$$C \leq 2(A^p + B^p)^{1/p} < \infty.$$

With this understood, it follows from the Hölder inequality that

$$\begin{aligned}
C^p &= \int_X f(f+g)^{p-1}d\mu + \int_X g(f+g)^{p-1}d\mu \\
&\leq \left(\int_X f^p d\mu\right)^{1/p}\left(\int_X (f+g)^{(p-1)q}d\mu\right)^{1/q} + \left(\int_X g^p d\mu\right)^{1/p}\left(\int_X (f+g)^{(p-1)q}d\mu\right)^{1/q} \\
&= (A+B)\left(\int_X (f+g)^p d\mu\right)^{1/q} = (A+B)C^{p-1}.
\end{aligned}$$

It follows that $C \leq A+B$ and this proves Theorem 10.1. □

Definition 10.2 Let $(X, \mathcal{A}, \mu)$ be a measure space and let $1 \leq p < \infty$. Let f be a measurable function on X. The L^p-norm of f is the number

$$||f||_p = \left(\int_X |f|^p d\mu\right)^{1/p}. \tag{10.6}$$

A function $f : X \to \mathbb{R}$ is called p-integrable or an L^p-function if it is measurable and $||f||_p < \infty$. The space of L^p-functions is denoted by

$$\mathfrak{L}^p(\mu) := \{f : X \to \mathbb{R}|\ f\ is\ \mathcal{A} - measurable\ and\ ||f||_p < \infty\}. \tag{10.7}$$

It follows from the Minkowski inequality that the sum of two L^p-functions is again an L^p-function, and hence $\mathfrak{L}^p(\mu)$ is a real vector space. Moreover, the function

$$\mathfrak{L}^p(\mu) \to [0, \infty] : f \to ||f||_p$$

satisfies the triangle inequality

$$||f+g||_p \leq ||f||_p + ||g||_p,\ \ f, g \in \mathfrak{L}^p$$

and

$$||\lambda f||_p = |\lambda| ||f||_p,$$

for all $\lambda \in \mathbb{R}$ and $f \in \mathfrak{L}^p(\mu)$. However, in general $||\cdot||_p$ is not a norm on $\mathfrak{L}^p(\mu)$ because $||f||_p = 0$ if and only if $f = 0$ almost everywhere. We can turn the space $\mathfrak{L}^p(\mu)$ into a normed vector space by identifying two functions $f, g \in \mathfrak{L}^p(\mu)$ whenever they agree almost everywhere. Thus we introduce the equivalence relation

$$f \sim g \Leftrightarrow f = g\ \ \mu - almost\ everywhere. \tag{10.8}$$

Denote the equivalence class of a function $f \in \mathfrak{L}^p(\mu)$ under this equivalence relation by $[f]$ and the quotient space by

$$L^p(\mu) = \mathfrak{L}^p(\mu)/\sim.$$

This is again a real vector space. The L^p-norm in (10.6) depends only on the equivalence class of f and so the map

$$L^p(\mu) \to [0, \infty) : [f] \to ||f||_p$$

is well defined. It is a norm on $L^p(\mu)$. Thus we have defined the normed vector space $L^p(\mu)$ for $1 \leq p < \infty$. It is often convenient to abuse notation and write

$f \in L^p(\mu)$ instead of $[f] \in L^p(\mu)$, always bearing in mind that then f denotes an equivalence class of p-integrable functions.

Remark 10.3 Assume $1 < p < \infty$ and let $f, g \in L^p(\mu)$ such that

$$||f+g||_p = ||f||_p + ||g||_p, \ ||f||_p \neq 0.$$

Then there exists a real number $\lambda \geq 0$ such that $g = \lambda f$ almost everywhere (see Exercises 9, 1(ii)).

Example 10.4 If $\mu : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ is the counting measure we write

$$l^p = L^p(\mu).$$

If

$$f : \mathbb{N} \to \mathbb{R}$$

is defined by

$$f(n) := x_n \ for \ n \in \mathbb{N},$$

then

$$\int_{\mathbb{N}} |f|^p d\mu = \sum_{n=1}^{\infty} |x_n|^p.$$

Thus the elements of l^p are sequences $(x_n)_{n=1}^{\infty}$ of real numbers such that

$$||(x_n)||_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < \infty.$$

Next, we define $L^\infty(\mu)$.

Definition 10.5 Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f : X \to [0, \infty]$ be a measurable function and set

$$G(f) = \{c : \mu(\{x : |f(x)| > c\}) = 0\},$$

and define the *essential supremum* of f, denoted by $||f||_\infty$, as

$$||f||_\infty = \text{ess sup} f = \begin{cases} \inf G(f), & \text{if } G(f) \neq \emptyset, \\ \infty, & \text{if } G(f) = \emptyset. \end{cases} \tag{10.9}$$

A function $f : X \to \mathbb{R}$ is called an L^∞-function if it is measurable and

$$||f||_\infty < \infty. \tag{10.10}$$

The set of L^∞-functions on X will be denoted by

$$\mathfrak{L}^\infty(\mu) := \{f : X \to \mathbb{R} | \ f \text{ is } \mathcal{A} - \text{measurable and } ||f||_\infty < \infty\}.$$

and the quotient space by the equivalence relation (10.8) by

$$L^\infty(\mu) = \mathfrak{L}^\infty(\mu)/\sim. \tag{10.11}$$

This is a normed vector space with the norm defined by (10.9), which depends only on the equivalence class of f.

Lemma 10.6 *Let $f \in L^\infty(\mu)$.*
a) If $\alpha \in G(f)$ then $[\alpha, \infty) \subset G(f)$.
b) $||f||_\infty \in G(f)$.
c)

$$\mu(\{x \in X | \ |f(x)| > \alpha\}) = 0 \Leftrightarrow \alpha \geq ||f||_\infty$$

and

$$\mu(\{x \in X|\ |f(x)| > \alpha\}) > 0 \Leftrightarrow \alpha < ||f||_\infty.$$

d) $|f| \leq ||f||_\infty$ *a.e.*

Proof.
a) For any $\beta \in [\alpha, \infty)$, $\{x : |f(x)| > \beta\} \subset \{x : |f(x)| > \alpha\}$.
b) Let $t_1 > t_2 > \cdots \in G(f)$ such that $t_n \to ||f||_\infty$ as $n \to \infty$. Then by the continuity from below of μ

$$\mu(\{x \in X|\ |f(x)| > ||f||_\infty\}) = \lim_{n\to\infty} \mu(\{x \in X|\ |f(x)| > t_n\}) = 0.$$

c) By a) and b), if $\alpha \geq ||f||_\infty$, then $\alpha \in G(f)$, and by the definition of $||f||_\infty$, if $\alpha < ||f||_\infty$, then $\alpha \notin G(f)$.
d) Because of b). □

Notice also that $||f||_\infty$ is characterized by the property

$$||f||_\infty \leq M \Leftrightarrow |f| \leq M \text{ a.e.}$$

Observe that $f \in \mathcal{L}^\infty$ iff there is a bounded measurable function g such that $f = g$ a.e.. In fact, we can take $g = f\chi_A$ where $A = \{x : |f(x)| \leq ||f||_\infty\}$.

If $\{x_n\}$ is a sequence in a normed vector space V, the series $\sum_{n=1}^\infty x_n$ is said to converge to x if $\sum_{n=1}^N x_n \to x$ as $N \to \infty$, and it is called absolutely convergent if $\sum_{n=1}^\infty ||x_n|| < \infty$

Lemma 10.7 *A normed vector space V is complete iff every absolutely convergent series in V converges.*

Proof. If V is complete and $\sum_{n=1}^\infty ||x_n|| < \infty$, let $S_N = \sum_{n=1}^N x_n$. Then for $N > M$, we have

$$||S_N - S_M|| \leq \sum_{n=M+1}^{N} ||x_n|| \to 0, \text{ as } M, N \to \infty.$$

so the sequence $\{S_N\}$ is Cauchy and hence convergent. Conversely, suppose that every absolutely convergent series converges, and let $\{x_n\}$ be a Cauchy sequence. We can choose $n_1 < n_2 < \cdots$ such that $||x_n - x_m|| < 2^{-j}$ for $m, n \geq n_j$. Let $y_1 = x_{n_1}$, $y_j = x_{n_j} - x_{n_{j-1}}$ for $j > 1$. Then $\sum_{j=1}^k y_j = x_{n_k}$, and

$$\sum_{j=1}^\infty ||y_j|| \leq ||y_1|| + \sum_{j=1}^\infty \frac{1}{2^j} < \infty,$$

so $\lim_{k\to\infty} x_{n_k} = \sum_{j=1}^\infty y_j$ exists. But since $\{x_n\}$ is Cauchy, we know that $\{x_n\}$ converges to the same limit as $\{x_{n_k}\}$. □

Theorem 10.8 $L^p(\mu)$ *is a Banach space for* $1 \leq p \leq \infty$.

Proof. We consider firstly the case $1 \le p < \infty$. We use Lemma 10.7. Suppose $\{f_i\} \subset L^p$ and $\sum_{i=1}^\infty ||f_i||_p = B < \infty$. Let $G_n = \sum_{i=1}^n |f_i|$ and $G = \sum_{i=1}^\infty |f_i|$. Then $||G_n||_p \le \sum_{i=1}^n ||f_i||_p \le B < \infty$ for all n, and by the monotone convergence theorem,

$$\int_X G^p d\mu = \lim_{n\to\infty} \int_X G_n^p d\mu \le B^p.$$

Hence $G \in L^p$, and in particular $G(x) < \infty$ a.e., which implies that the series $\sum_{i=1}^\infty f_i$ converges a.e. Denoting its sum by F, we have $|F| \le G$ and hence $F \in L^p$; moreover, $|F - \sum_{i=1}^n f_i|^p \le (2G)^p \in L^1$, so by the dominated convergence theorem,

$$||F - \sum_{i=1}^n f_i||_p^p = \int_X \left| F - \sum_{i=1}^n f_i \right|^p d\mu \to 0.$$

Thus the series $\sum_{i=1}^\infty f_i$ converges to F in the L^p norm.

Suppose now that $p = +\infty$ and that $\{f_k\}$ is a sequence of functions that belong to $L^\infty(\mu)$ and satisfy $\sum_{k=1}^\infty ||f_k||_\infty < \infty$. For each positive integer k, let

$$E_k = \{x \in X : |f_k(x)| > ||f_k||_\infty\}, \quad E = \cup_{k=1}^\infty E_k.$$

We have $\mu(E_k) = 0$ and so $\mu(E) = 0$. Hence

$$\sum_{k=1}^\infty |f_k(x)| \le \sum_{k=1}^\infty ||f_k||_\infty < \infty$$

for all $x \in E^c$, which implies that $\sum_{k=1}^\infty f_k(x)$ converges for all $x \in E^c$. Define

$$f(x) = \begin{cases} \sum_{k=1}^\infty f_k(x), & \text{if } x \notin E, \\ 0, & \text{if } x \in E. \end{cases}$$

Then

$$\left| f(x) - \sum_{k=1}^n f_k(x) \right| = \left| \sum_{k=n+1}^\infty f_k(x) \right| \le \sum_{k=n+1}^\infty |f_k(x)| \le \sum_{k=n+1}^\infty ||f_k||_\infty \ a.e.$$

Thus

$$||f - \sum_{k=1}^n f_k||_\infty \le \sum_{k=n+1}^\infty ||f_k||_\infty.$$

Letting $n \to \infty$, we have $\sum_{k=n+1}^\infty ||f_k||_\infty \to 0$, so $||f - \sum_{k=1}^n f_k||_\infty \to 0$, i.e., $\sum_{k=1}^n f_k \to f$ in L^∞. Therefore, L^∞ is a Banach space.

□

Proposition 10.9 *i) $f_n \to f$ in $L^\infty \Leftrightarrow f_n \to f$ uniformly a.e..*

ii) For any $1 \le p \le \infty$, simple measurable functions determine a dense subspace of L^p.

Proof. i) Assume $||f_n - f||_\infty \to 0$ as $n \to \infty$. Let

$$E_n = \{x : |f_n(x) - f(x)| > ||f_n - f||_\infty\}, \ E = \cup_{n=1}^\infty E_n.$$

Then $\mu(E_n) = 0$ and hence $\mu(E) = 0$. For all $x \in E^c$, $|f_n(x) - f(x)| \le ||f_n - f||_\infty$ for all n. So $f_n \to f$ uniformly on E^c. Conversely, if $f_n \to f$ uniformly on E^c for an E with $\mu(E) = 0$, then for any $\epsilon > 0$, there is an N such that $|f_n - f| < \epsilon$ a.e. for all $n \ge N$. Thus $||f_n - f||_\infty \le \epsilon$ for all $n \ge N$.

ii) Consider first the case that $1 \leq p < \infty$. Let $f \in L^p$. Since f is measurable, there is a sequence of simple measurable functions $\{s_n\}$ such that $|s_n| \leq |f|$ and $s_n \to f$ pointwise as $n \to \infty$. Since $|s_n - f|^p \leq 2^p|f|^p$ for all n and $2^p|f|^p \in L^1$, by Dominant Convergence Theorem

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

i.e.,

$$||s_n - f||_p \to 0.$$

Assume now that $p = \infty$. Let $f \in L^\infty$ and define

$$E = \{x : |f(x)| > ||f||_\infty\}.$$

Then $\mu(E) = 0$. Since f is bounded on E^c, it follows from Theorem 7.3 that there is a sequence of simple functions $\{f_n\}$ which converges to f uniformly on E^c. Then by (i), $||f_n - f||_\infty \to 0$.

□

10.1. Dual of $L^p(\mu)$. When studying normed linear spaces, one of the important objectives is to identify the dual space of a normed linear space, i.e. the space of all continuous linear functionals on that space. In this section, we will identify the dual spaces of $L^p(\mu)$.

Let $(X, \mathcal{A}, \mu)$ be a measure space and let $1 \leq p \leq \infty$. Let q be the conjugate exponent of p. Let $g \in L^q(\mu)$ and define $T_g : L^p(\mu) \to \mathbb{R}$ by

$$T_g(f) = \int_X fgd\mu, \;\; f \in L^p(\mu).$$

T_g is a linear functional defined on $L^p(\mu)$ and is also continuous. In fact, by Hölder's inequality, we have

$$|T_g(f)| \leq ||f||_p||g||_q,$$

which shows the continuity of T_g. We also have

$$||T_g|| \leq ||g||_q. \tag{10.12}$$

In this section, we will show that if $(X, \mathcal{A}, \mu)$ is a σ-finite measure space, then every continuous linear functional on $L^p(\mu)$ occurs in this way, if $1 \leq p < \infty$ and that we have equality in (10.12). Thus, we have an isometric isomorphism (i.e. an isomorphism which preserves norms) between the dual of $L^p(\mu)$ and the space $L^q(\mu)$ and so we can identify the latter space with the dual of the former, when $1 \leq p < \infty$.

Proposition 10.10 (Uniqueness) *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. Let $1 \leq p < \infty$. If $g_i, i = 1, 2$, are in $L^q(\mu)$, where q is the conjugate exponent of p, such that $T_{g_1} = T_{g_2}$, then $g_1 = g_2$ a.e.*

Proof. If $f \in L^p(\mu)$, we have

$$\int_X f(g_1 - g_2)d\mu = 0.$$

Let $E \subset X$ be a measurable set of finite measure; then $\chi_E \in L^p(\mu)$ and so we have

$$\int_E (g_1 - g_2) d\mu = 0.$$

Since $E^+ = \{x \in E : (g_1 - g_2)(x) > 0\}$ is also measurable with finite measure, we conclude that

$$\int_{E^+} (g_1 - g_2) d\mu = 0,$$

which gives $\mu(E^+) = 0$. Similarly, one has $\mu(\{x \in E : (g_1 - g_2)(x) < 0\}) = 0$. Consequently, $g_1 - g_2 = 0$ on E a.e. By the σ-finiteness, we can find a collection of disjoint sets $\{E_i\}_{i=1}^\infty$ in $\mathcal{A}$ such that $X = \overset{\infty}{\underset{i=1}{\cup}} E_i$ and $\mu(E_i) < \infty$ for each i. It follows from $g_1 - g_2 = 0$ on E_i a.e. for each i that $g_1 = g_2$ almost everywhere. $\square$

Thus the mapping $g \to T_g$ from $L^q(\mu)$ into the dual of $L^p(\mu)$ is injective and continuous by virtue of (10.12). We need to show that it is surjective and that it is an isometry. We will first prove this when the measure space is finite and then deduce the general case of σ-finite measure spaces.

The proof of the following fact is left as an exercise.

Lemma 10.11 *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and let $g : X \to \mathbb{R}$ be a measurable function such that*

$$\left| \frac{1}{\mu(E)} \int_E g d\mu \right| \leq K, \tag{10.13}$$

for all $E \in \mathcal{A}$ with $\mu(E) > 0$. Then $|g| \leq K$ almost everywhere.

Theorem 10.12 *Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Let $1 \leq p < \infty$. Let T be a continuous linear functional on $L^p(\mu)$. Then, there exists a unique $g \in L^q(\mu)$ such that $T = T_g$ and $||T|| = ||g||_q$.*

Proof. **Step 1.** For $E \in \mathcal{A}$, define

$$\lambda(E) = T(\chi_E).$$

This is well-defined, since $\mu(E) < \infty$ and so $\chi_E \in L^p(\mu)$. If A and B are disjoint measurable sets, then

$$\chi_{A \cup B} = \chi_A + \chi_B,$$

and so, by the linearity of T, we conclude that λ is finitely additive. Let $\{E_i\}_{i=1}^\infty$ be a countable collection of disjoint sets in $\mathcal{A}$. Let $E = \overset{\infty}{\underset{i=1}{\cup}} E_i$ and set $F_k = \cup_{i=1}^k E_i$. Then $F_k \uparrow E$ and by using the finiteness of $\mu(E)$ we have

$$\mu(E - F_k) \to 0.$$

Now,

$$||\chi_E - \chi_{F_k}||_p = (\mu(E - F_k))^{1/p},$$

and so $\chi_{F_k} \to \chi_E$ in $L^p(\mu)$. Then $T(\chi_{F_k}) \to T(\chi_E)$ which shows that

$$\lambda(E) = \sum_{i=1}^\infty \lambda(E_i).$$

Thus λ defines a signed measure on $(X, \mathcal{A})$. Furthermore, if $\mu(E) = 0$, then $\chi_E = 0$ in $L^p(\mu)$ and so $\lambda(E) = 0$ as well. Thus $\lambda \ll \mu$. Hence, by the Radon-Nikodym theorem, there exists a integrable function g such that, for every $E \in \mathcal{A}$, we have

$$\lambda(E) = \int_E g d\mu,$$

or, in other words,

$$T(\chi_E) = \int_X \chi_E g d\mu.$$

Step 2. By linearlity, the above equality holds for simple measurable functions on X. Let $f \in L^\infty(\mu)$; then $f \in L^p(\mu)$. Since f is measurable, there is a sequence of simple measurable functions $\{s_n\}$ such that $|s_n| \leq |f|$ and $s_n \to f$ pointwise as $n \to \infty$. Then $s_n \to f$ in $L^p(\mu)$ and so $T(s_n) \to T(f)$. On the other hand, $s_n g \to fg$ pointwise and $|s_n g| \leq |f||g| \leq ||f||_\infty |g|$. Since g is integrable, it follows, from the dominated convergence theorem, that

$$\int_X s_n g d\mu \to \int_X f g d\mu.$$

Thus, for all bounded functions f, we have

$$T(f) = \int_X f g d\mu. \tag{10.14}$$

Step 3. Let $p = 1$. Let $E \in \mathcal{A}$ with $\mu(E) > 0$. Then

$$\left| \int_E g d\mu \right| = \left| \int_X g \chi_E d\mu \right| = |T(\chi_E)| \leq ||T|| \cdot ||\chi_E||_1 = ||T|| \cdot \mu(E).$$

Thus

$$\left| \frac{1}{\mu(E)} \int_E g d\mu \right| \leq ||T||.$$

It then follows from Lemma 10.11 that $|g| \leq ||T||$ almost everywhere. Thus, in this case $g \in L^\infty(\mu)$ and $||g||_\infty \leq ||T||$.

Step 4. Let $1 < p < \infty$. Let h be a measurable function (taking values ± 1) such that $hg = |g|$. Let

$$E_n = \{x \in X | |g(x)| \leq n\}, \ n \in \mathbb{N}.$$

Set $f = \chi_{E_n} |g|^{q-1} h$; then

$$|f|^p = \chi_{E_n} |g|^q.$$

Furthermore, by the definition of E_n, it follows that f is bounded. Since $fg = \chi_{E_n} |g|^q$, we get, from **Step 2**, that

$$\int_{E_n} |g|^q d\mu = \int_X f g d\mu = T(f),$$

and so

$$\int_{E_n} |g|^q d\mu \leq ||T|| \cdot ||f||_p = ||T|| \left(\int_{E_n} |g|^q d\mu \right)^{1/p},$$

which yields

$$\left(\int_{E_n} |g|^q d\mu \right)^{1/q} \leq ||T||.$$

But $E_n \uparrow X$ and so, by the monotone convergence theorem, we have

$$\left(\int_X |g|^q d\mu \right)^{1/q} \leq ||T||.$$

Thus $g \in L^q(\mu)$ and $||g||_q \leq ||T||$.

Step 5. Let $1 \leq p < \infty$. Then, we have seen that $g \in L^q(\mu)$ and that $||g||_q \leq ||T||$. Observe that simple measurable functions form a dense subspace $\mathbb{S}$ of $L^p(\mu)$. Both sides of (10.14) define continuous linear functionals on $L^p(\mu)$ and agree on $\mathbb{S}$ and so, they agree on all of $L^p(\mu)$. Thus, we get that, in fact, $T = T_g$, in which case, we have that

$$||T|| = ||T_g|| \leq ||g||_q \leq ||T||.$$

Thus $T = T_g$ and $||T|| = ||T_g|| = ||g||_q$. □

Theorem 10.13 (Riesz representation theorem) *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space. Let $1 \leq p < \infty$. Let T be a continuous linear functional on $L^p(\mu)$. Then, there exists a unique $g \in L^q(\mu)$ such that $T(f) = \int_X fg d\mu$ for all $f \in L^p$, and $||T|| = ||g||_q$.*

Proof. Let $\{E_n\}$ be an increasing sequence of sets such that $0 < \mu(E_n) < \infty$ and $X = \cup_{n=1}^\infty E_n$, and let us agree to identify $L^p(E_n)$ and $L^q(E_n)$ with the subspaces of $L^p(X)$ and $L^q(X)$ consisting of functions that vanish outside E_n. According to Theorem 10.12 for each n there exists $g_n \in L^q(E_n)$ such that $T(f) = \int_{E_n} fg_n$ for all $f \in L^p(E_n)$, and $||g_n||_q = ||T|_{L^p(E_n)}|| \leq ||T||$. The function g_n is unique modulo alterations on sets of measure zero, so $g_n = g_m$ a.e. on E_n for $n < m$, and we can define g a.e. on X by setting $g = g_n$ on E_n. By the monotone convergence theorem, $||g||_q = \lim_{n\to\infty} ||g_n||_q \leq ||T||$, so $g \in L^q$. Moreover, if $f \in L^p$, then by the dominated convergence theorem, $f\chi_{E_n} \to f$ in the L^p norm and hence $T(f) = \lim_{n\to\infty} T(f\chi_{E_n}) = \lim_{n\to\infty} \int_{E_n} fg = \int_X fg$. Also, it follows from the proof of Theorem 10.12 that $||T|| = ||g||_q$. □

Corollary *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space with $1 < p < \infty$; then $L^p(\mu)$ is reflexive.*

The Riesz Weak Compactness Theorem *Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and $1 < p < \infty$. Then every bounded sequence in $L^p(\mu)$ has a weakly convergent subsequence; that is, if $\{f_n\}$ is a bounded sequence $L^p(\mu)$, then there is a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ and a function f in $L^p(\mu)$ for which*

$$\lim_{k\to\infty} \int_X f_{n_k} g d\mu = \int_X fg d\mu, \ \forall g \in L^q(\mu), \text{ where } \frac{1}{p} + \frac{1}{q} = 1.$$

Proof. The preceding proposition asserts that $L^p(\mu)$ is reflexive. However, every bounded sequence in a reflexive Banach space has a weakly convergent subsequence. The conclusion now follows from the Riesz Representation Theorem for the dual of $L^p(\mu)$.

Proposition 10.15 *If $1 < p < q < r \leq \infty$, then $L^q \subset L^p + L^r$; that is, each $f \in L^q$ is the sum of a function in L^p and a function in L^r.*

Proof. If $f \in L^q$, let $E = \{x : |f(x)| > 1\}$ and set $g = f\chi_E$ and $h = f\chi_{E^c}$. Then $|g|^p = |f|^p\chi_E \leq |f|^q\chi_E$, so $g \in L^p$, and $|h|^r = |f|^r\chi_{E^c} \leq |f|^q\chi_{E^c}$, so $h \in L^r$. (For $r = \infty$, obviously $||h||_\infty \leq 1$.) $\square$

Proposition 10.16 *If $1 < p < q < r \leq \infty$, then $L^p \cap L^r \subset L^q$ and for $f \in L^p \cap L^r$, we have $||f||_q \leq ||f||_p^\lambda ||f||_r^{1-\lambda}$, where $\lambda \in (0,1)$ is given by*

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}.$$

Proof. If $r = \infty$, then

$$|f|^q \leq ||f||_\infty^{q-p}|f|^p,\ \lambda = \frac{p}{q}.$$

Thus

$$||f||_q \leq ||f||_\infty^{1-\frac{p}{q}}||f||_p^{\frac{p}{q}} = ||f||_\infty^{1-\lambda}||f||_p^\lambda.$$

If $r < \infty$, we use Hölder's inequality, taking the pair of conjugate exponents to be $\frac{p}{\lambda q}$ and $\frac{r}{(1-\lambda)q}$:

$$\int |f|^q = \int |f|^{\lambda q} \cdot |f|^{(1-\lambda)q} \leq |||f|^{\lambda q}||_{\frac{p}{\lambda q}} \cdot |||f|^{(1-\lambda)q}||_{\frac{r}{(1-\lambda)q}} = ||f||_p^{\lambda q} \cdot ||f||_r^{(1-\lambda)q}.$$

$\square$

Exercises 9

1. (i) Assume $0 < \int_X f^p d\mu < \infty$ and $0 < \int_X g^q d\mu < \infty$ Prove that equality holds in (10.3) if and only if there exists a constant $\alpha > 0$ such that $g^q = \alpha f^p$ almost everywhere. *Hint.* Use the proof of the Hölder inequality and the fact that equality holds in (9.2) if and only $p^p = b^q$.

(ii) Assume $0 < \int_X f^p d\mu < \infty$ and $0 < \int_X g^q d\mu < \infty$ Prove that equality holds in (9.4) if and only if there exists a constant $\beta > 0$ such that $g = \beta f$ almost everywhere. *Hint.* Use part (i) and the proof of the Minkowski inequality.

2. Suppose X is a metric space in which every Cauchy sequence has a convergent subsequence. Show that X is a complete metric space.

3. Let $f \in \mathfrak{L}^\infty(\mu)$. Show that there exists a measurable set $E \subset \mathcal{A}$ such that

$$\mu(E) = 0 \textit{ and } \sup_{X-E} |f| = ||f||_\infty.$$

4. Prove Lemma 10.11.

6. Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 < p < \infty$. Prove that $f \in L^p$ iff

$$\sum_{n=1}^{\infty} n^{-(p+1)} \mu\{|f| > 1/n\} + \sum_{n=1}^{\infty} n^{p-1} \mu\{|f| > n\} < \infty.$$

7. Let $(X, \mathcal{A}, \mu)$ be a finite measure space $1 \leq p < q < \infty$, and $a > 0$. Prove that

$$||f||_p \leq a\mu(\{|f| \leq a\})^{1/p} + ||f||_q \mu(\{|f| > a\})^{1/p-1/q}.$$

8. Let $\{f_n\} \subset L^1(\mu)$ be nonnegative and satisfying $\liminf_n f_n \geq f$ a.e. in X. Show that

$$\int_X f_n d\mu = \int_X f d\mu = 1 \Rightarrow \lim_n \int_X |f_n - f| d\mu = 0.$$

Hint: notice that the positive part and the negative part of $f - f_n$ have the same integral to obtain

$$\int_X |f_n - f| d\mu = 2 \int_X (f_n - f)^+ d\mu.$$

Then, apply the dominated convergence theorem.

9. Show that the following extension of Fatou's lemma: if $f_n \geq -g_n$, with $g_n \in L^1$ nonnegative, $g_n \to g$ in L^1, then

$$\liminf_{n\to\infty} \int_X f_n d\mu \geq \int_X \liminf_{n\to\infty} f_n d\mu,$$

whenever the right hand side is well defined.

Hint: prove first the statement under the additional assumption that $g_n \to g$ a.e. in X.

10. Let $\{f_n\} \subset L^1(\mu)$ be nonnegative functions. Show that the conditions

$$\liminf_{n\to\infty} f_n \geq f \ a.e., \quad \limsup_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu < \infty$$

imply the convergence of f_n to f in L^1. Hint: use exercise 8.

11. Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Show that $L^\infty \subset L^p$ for all $p \in (0, \infty)$. Also, show that, if f is measurable, then

$$\lim_{p\to\infty} ||f||_p = ||f||_\infty.$$

12. Let $0 < p < 1$ and f, g be measurable functions on X such that $f \geq 0, g \geq 0$. Then

$$||f + g||_p \geq ||f||_p + ||g||_p.$$

(This is Minkowski's Inequality for $0 < p < 1$.)

13. a) Let $p \geq 1$ and let $||f_n - f||_p \to 0$ as $n \to \infty$. Show that $||f_n||_p \to ||f||_p$ as $n \to \infty$.

b) Suppose $f_n, f \in L^p$, $f_n \to f$ a.e and $||f_n||_p \to ||f||_p$ as $n \to \infty$. Show that $||f_n - f||_p \to \infty$ as $n \to \infty$.

14. Show that, if $k_1, k_2, \cdots, k_n > 1, \sum_{i=1}^{n} \frac{1}{k_i} = 1$, and $f_i \in L^{k_i}$ for each i, then

$$\int_X |f_1 \cdots f_n| d\mu \leq ||f_1||_{k_1} \cdots ||f_n||_{k_n}.$$

15. Let p and q be conjugate indices and $\lim_{n\to\infty} ||f_n - f||_p = 0 = \lim_{n\to\infty} ||g_n - g||_q$, where $f_n, f \in L^p$ and $g_n, g \in L^q, n = 1, 2, \cdots$. Show that $\lim_{n\to\infty} ||f_n g_n - fg||_1 = 0$.

16. Let $f_n, f \in L^p$ and $||f_n - f||_p \to 0$, $f_n \to g$ a.e. as $n \to \infty$. What relation exists between f and g.

17. Let $p, q \in (1, \infty)$. Show that the inclusion $L^p \subset L^q$ implies that the inclusion map $i : L^p \to L^q$ is continuous.

Hint: use closed graph theorem.

18. Let $(X, \mathcal{A}, \mu)$ be a finite measure space. Assume that $f \in L^\infty, ||f||_\infty > 0$. Let

$$a_n = \int_X |f|^n d\mu, \ n = 1, \cdots.$$

Prove that

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = ||f||_\infty.$$

19. Let $(X, \mathcal{A}, \mu)$ be a measure space. Assume that $f, g \in L^p$.

i) If $0 < p < 1$, prove that

$$\int_X ||f|^p - |g|^p| \, d\mu \leq \int_X |f - g|^p| \, d\mu$$

that $d(f, g) = \int_X |f - g|^p| \, d\mu$ define a metric on L^p, and that the resulting metric is complete.

iii) If $1 \leq p < \infty$ and $||f||_p \leq R, ||g||_p \leq R$, prove that

$$\int_X ||f|^p - |g|^p| \, d\mu \leq 2pR^{p-1}||f - g||_p.$$

Hint: Prove first, for $x \geq 0, y \geq 0$, that

$$|x^p - y^p| \leq \begin{cases} |x - y|^p, & if \ 0 < p < 1, \\ p|x - y|(x^{p-1} + y^{p-1}), & if \ 1 \leq p < \infty. \end{cases}$$

20. Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $f : X \to (0, \infty)$ satisfies $\int_X f d\mu = 1$. Prove that for every $E \in \mathcal{A}$ with $0 < \mu(E) < \infty$, that

$$\int_E (\log f) d\mu \leq \mu(E) \log \frac{1}{\mu(E)}$$

and, when $0 < p < 1$,

$$\int_E f^p d\mu \leq \mu(E)^{1-p}.$$

21. Let $(X, \mathcal{A}, \mu)$ be a measure space, $2 < p < \infty$, and $\{f_n\}, f$ defined on X with $||f_n||_p \to ||f||_p$, such that $f_n \rightharpoonup f$ in L^p. Prove that $f_n \to f$ in L^p.

22. Let $f \in L^p \cap L^\infty$ for some $p < \infty$, so that $f \in L^q$ for all $q > p$. Prove: $||f||_\infty = \lim_{q\to\infty} ||f||_q$.

11. The Lebesgue measure on $\mathbb{R}^n$

The purpose of this section is to introduce the Lebesgue outer measure m^* on $\mathbb{R}^n$, construct the Lebesgue measure as the restriction of m^* to the σ-algebra of all m^*-measurable subsets of $\mathbb{R}^n$, and prove its basic properties.

Definition 11.1 A *closed cuboid* in $\mathbb{R}^n$ is a set of the form

$$\begin{aligned} Q &\equiv Q(a;b) \\ &:= [a_1, b_1] \times [a_2, b_2] \times \cdots [a_n, b_n] \\ &= \{x = (x_1, \cdots, x_n) \in \mathbb{R}^n : a_j \leq x_j \leq b_j \ for\ j = 1, \cdots, n\}. \end{aligned} \tag{11.1}$$

for $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{R}^n$ with $a_j < b_j$ for all j. If the edge lengths $b_k - a_k$ are all equal, Q will be called an n-dimensional cube with edges parallel to the coordinate axes. The $(n - dimensional)$ volume of the cuboid $Q(a;b)$ is defined by

$$|Q(a;b)| = \prod_{i=1}^{n} (b_i - a_i).$$

The volume of the open cuboid

$$Q^o := int(Q(a;b)) = (a_1, b_1) \times \cdots \times (a_n, b_n)$$

is defined by

$$|Q^o| = |Q(a;b)|.$$

A set equal to a closed cuboid minus some part of its boundary will be called a partly open cuboid. We know that every open set in $\mathbb{R}^1$ can be written as a countable union of disjoint open intervals.

In the following lemma we will be dealing with a certain collection of half-open cubes, namely with those that have the form

$$\{(x_1, \cdots, x_n) : j_i 2^{-k} \leq x_i < (j_i + 1)2^{-k} \text{ for } i = 1, \cdots, n\} \tag{11.2}$$

for some integers $j_1, ..., j_n$ and some positive integer k.

Lemma 11.2 *Each open subset of $\mathbb{R}^n$ is the union of a countable disjoint collection of half-open cubes, each of which is of the form given in expression (11.2).*

Proof. For each positive integer k let $\mathcal{C}_k$ be the collection of all cubes of the form

$$\{(x_1, \cdots, x_n) : j_i 2^{-k} \leq x_i < (j_i + 1)2^{-k} \text{ for } i = 1, \cdots, n\}.$$

where $j_1, ..., j_n$ are arbitrary integers. It is easy to see that

a) each $\mathcal{C}_k$ is a countable partition of $\mathbb{R}^n$, and
(b) if $k_1 < k_2$, then each cube in $\mathcal{C}_{k_2}$ is included in some cube in $\mathcal{C}_{k_1}$.

Let $\mathcal{C} = \cup_{k=1}^\infty \mathcal{C}_k$. Suppose that U is an open subset of $\mathbb{R}^n$. For any $x \in U$, let $C_x \in \mathcal{C}$ be the biggest cube such that $x \in C_x \subset U$. This cube is uniquely defined: indeed, fix an x; for any k there is only one cube $C_{x,k} \in \mathcal{C}_k$ such that $x \in C_{x,k}$; moreover, since U is open, for k large enough $C_{x,k} \subset U$; we can then define $C_x = C_{x,k_0}$ where k_0 is the smallest integer k such that $C_{x,k} \subset U$.

This family $\{C_x\}_{x\in U}$ is a partition of U, that is, for any $x, y \in U$, either $C_x = C_y$ or $C_x \cap C_y = \emptyset$; indeed, if we suppose that $C_x \cap C_y \neq \emptyset$, then one of the two cubes is contained in the other, say $C_x \subset C_y$. This leads to $x \in C_x \subset C_y \subset U$, contradicting to the definition of C_x unless $C_x = C_y$. □

Every open set in $\mathbb{R}^n, n \geq 1$, can be written as a countable union of nonoverlapping (closed) cubes. It can also be written as a countable union of disjoint half open cubes.

The set of all closed cuboids in $\mathbb{R}^n$ will be denoted by

$$\mathcal{Q}_n = \{Q(a;b) : a_1, \cdots, a_n; b_1, \cdots, b_n \in \mathbb{R}; a_j < b_j, j = 1, \cdots, n\}.$$

Definition 11.2 A subset $A \subset \mathbb{R}^n$ is called a Lebesgue null set if, for every $\epsilon > 0$, there exist a sequence of closed cuboids $Q_i \in \mathcal{Q}_n, i \in \mathbb{N}$ such that

$$A \subset \cup_{i=1}^{\infty} Q_i, \ \sum_{i=1}^{\infty} |Q_i| < \epsilon.$$

Definition 11.3 The Lebesgue outer measure on $\mathbb{R}^n$ is the function

$$\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$$

defined by

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^{\infty} |Q_i|, Q_i \in \mathcal{Q}_n, i = 1, \cdots, A \subset \cup_{i=1}^{\infty} Q_i \right\}. \tag{11.3}$$

We need the Carathéodory criterion to study the Lebesgue outer measure μ^*.

Theorem 11.4 (Carathéodory criterion). *Let (X, d) be a metric space. Consider an outer measure $\mu^* : \mathcal{P}(X) \to [0, \infty]$. Let $\mathcal{A}(\mu^*)$ be the σ-algebra consisting of μ^*-measurable subsets of X and $\mathcal{B}(X) \subset \mathcal{P}(X)$ be the Borel σ-algebra of (X, d). Then the following statements are equivalent:*

(i) $\mathcal{B}(X) \subset \mathcal{A}(\mu^*)$;

(ii) if $A, B \subset X$ satisfy $d(A, B) := \inf_{a\in A, b\in B} d(a, b) > 0$; then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Proof. $(i) \Rightarrow (ii)$. Let $A, B \subset X$ such that $\epsilon := d(A, B) > 0$. Define

$$U = \cup_{a\in A} B_\epsilon(a),$$

where $B_\epsilon(a) = \{x \in X : d(x, a) < \epsilon\}$ is the open ball of radius ϵ centered at a. Then U is open, $A \subset U$, $U \cap B = \emptyset$. Hence $U \in \mathcal{B}(X) \subset \mathcal{A}(\mu^*)$ by assumption and hence $\mu^*(A \cup B) = \mu^*((A \cup B) \cap U) + \mu^*((A \cup B) - U) = \mu^*(A) + \mu^*(B)$. Thus the outer measure μ^* satisfies (ii).

$(ii) \Rightarrow (i)$. We show that every closed set $A \subset X$ is μ^*-measurable, i.e., $\mu^*(D) = \mu^*(D \cap A) + \mu^*(D - A)$ for all $D \subset X$. It suffices to prove the following.

Claim 1. Fix a closed set $A \subset X$ and a set $D \subset X$ such that $\mu^*(D) < \infty$. Then

$$\mu^*(D) \geq \mu^*(D \cap A) + \mu^*(D - A).$$

To see this, replace the set $D-A$ by the smaller set $D-U_k$, where

$$U_k = \cup_{a\in A} B_{\frac{1}{k}}(a).$$

For each $k\in\mathbb{N}$, U_k is open and $d(x,y)\geq \frac{1}{k}$ for all $x\in D\cap A$ and all $y\in (D-U_k)$. Hence

$$d(D\cap A, D-U_k)\geq \frac{1}{k}.$$

By (ii) and the property of μ^*, we have

$$\begin{aligned}\mu^*(D\cap A)+\mu^*(D-U_k) &= \mu^*((D\cap A)\cup(D-U_k))\\ &\leq \mu^*((D\cap A)\cup(D-A))=\mu^*(D). \qquad (11.4)\end{aligned}$$

Now we prove

$$\lim_{k\to\infty}\mu^*(D-U_k)=\mu^*(D-A) \qquad (11.5)$$

which, combining with (11.4), will imply **Claim 1**. Note that

$$A=\cap_{i=1}^{\infty}U_i$$

because A is closed. (If $x\in U_i$ for all $i\in\mathbb{N}$, then there exists a sequence $a_i\in A$ such that $d(a_i,x)<1/i$ and hence $x=\lim_{i\to\infty}a_i\in A$.) This implies

$$\begin{aligned}U_k-A &= \cup_{i=1}^{\infty}(U_k-U_i)=\cup_{i=k+1}^{\infty}(U_k-U_i)\\ &= \cup_{i=k}^{\infty}(U_k-U_{i+1})=\cup_{i=k}^{\infty}(U_i-U_{i+1}).\end{aligned}$$

Thus

$$\begin{aligned}D-A &= (D-U_k)\cup(D\cap(U_k-A))\\ &= (D-U_k)\cup(\cup_{i=k}^{\infty}(D\cap(U_i-U_{i+1})))\\ &= (D-U_k)\cup(\cup_{i=k}^{\infty}E_i), \qquad (11.6)\end{aligned}$$

where $E_i=D\cap(U_i-U_{i+1})=D\cap U_i-U_{i+1}$.

Claim 2. The outer measures of the E_i's satisfy

$$\sum_{i=1}^{\infty}\mu^*(E_i)<\infty.$$

Observe that **Claim 2** implies (11.5). In fact, in this case,

$$\epsilon_k:=\sum_{i=k}^{\infty}\mu^*(E_i)\to 0.$$

Moreover, it follows from equation (11.6) that

$$\mu^*(D-A)\leq\mu^*(D-U_k)+\sum_{i=k}^{\infty}\mu^*(E_i)=\mu^*(D-U_k)+\epsilon_k. \qquad (11.7)$$

Hence

$$\mu^*(D-A)-\epsilon_k\leq\mu^*(D-U_k)\leq\mu^*(D-A).$$

for every $k\in\mathbb{N}$. Since $\epsilon_k\to 0$, this implies (11.5).

The proof of Claim 2 relies on the next assertion.

Claim 3. $d(E_i,E_j)>0$ for $i\geq j+2$.

Claim 3 $\Rightarrow$ **Claim 2**. In fact, we have if **Claim 3** holds that

$$\sum_{i=1}^{n} \mu^*(E_{2i}) = \mu^*(\cup_{i=1}^{n} E_{2i}) \leq \mu^*(D)$$

and

$$\sum_{i=1}^{n} \mu^*(E_{2i-1}) = \mu^*(\cup_{i=1}^{n} E_{2i-1}) \leq \mu^*(D)$$

for every $n \in \mathbb{N}$. Hence

$$\sum_{i=1}^{\infty} \mu^*(E_i) \leq 2\mu^*(D) < \infty$$

and this shows that **Claim 3** implies **Claim 2**.

Proof of **Claim 3**. We show that

$$d(E_i, E_j) \geq \frac{1}{(i+1)(i+2)} \; for \; j \geq i+2.$$

To see this, fix indices i, j with $j \geq i+2$. Let $x \in E_i$ and $y \in X$ such that

$$d(x, y) < \frac{1}{(i+1)(i+2)}$$

Then $x \notin U_{i+1}$ because $E_i \cap U_{i+1} = \emptyset$. Hence

$$d(a, x) \geq \frac{1}{i+1} \; for \; all \; a \in A.$$

This implies

$$d(a, y) \geq d(a, x) - d(x, y) > \frac{1}{i+1} - \frac{1}{(i+1)(i+2)} = \frac{1}{i+2} \geq \frac{1}{j}$$

for all $a \in A$. Hence $y \notin U_j$ and consequently $y \notin E_j$, because $E_j \subset U_j$. This proves **Claim 3** and Theorem 11.4.

Theorem 11.5 *Let $\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ be the function defined by (11.3). Then the following holds.*
(i) μ^* is an outer measure;
(ii) μ^* is translation-invariant, i.e., for all $A \subset \mathbb{R}^n$ and all $x \in \mathbb{R}^n$

$$\mu^*(A + x) = \mu^*(A);$$

(iii) if $A, B \subset \mathbb{R}^n$ are such that $d(A, B) > 0$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$;
(iv) $\mu^*(Q^o) = \mu^*(Q) = |Q|$ for all $Q \in \mathcal{Q}_n$.

Proof. The empty set is contained in every cuboid $Q \in \mathcal{Q}_n$. Since there are cuboids with arbitrarily small volume, it follows that $\mu^*(\emptyset) = 0$. It follows from Proposition 5.3 that μ^* is an outer measure on $\mathbb{R}^n$ since there are $D_1, D_2 \cdots$, in $\mathcal{Q}_n$ such that $\mathbb{R}^n = \cup_{i=1}^{\infty} D_i$. This proves (i).

We prove (ii). If $A \subset \bigcup\limits_{i=1}^{\infty} Q_i$ with $Q_i \in \mathcal{Q}_n$, then $A + x \subset \bigcup\limits_{i=1}^{\infty} (Q_i + x)$ for every $x \in \mathbb{R}^n$ and $|Q_i + x| = |Q_i|$, by the defnition of the volume. Hence part (ii) follows from Definition 11.3.

We prove (iii). Let $A, B \subset \mathbb{R}^n$ are such that $d(A,B) > 0$. Choose a sequence of closed cuboids $Q_i \in \mathcal{Q}_n$ such that

$$A \cup B \subset \bigcup_{i=1}^{\infty} Q_i, \ \sum_{i=1}^{\infty} |Q_i| < \mu^*(A \cup B) + \epsilon.$$

Subdividing each Q_i into finitely many smaller cuboids, if necessary, we may assume without loss of generality that

$$\operatorname{diam}(Q_i) := \sup_{x,y \in Q_i} |x - y| < \frac{d(A,B)}{2}.$$

Here $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. Then, for every $i \in \mathbb{N}$, we have either $Q_i \cap A = \emptyset$ or $Q_i \cap B = \emptyset$. This implies

$$I \cap J = \emptyset, \ I := \{i \in \mathbb{N}, Q_i \cap A \neq \emptyset\}, \ J := \{i \in \mathbb{N}, Q_i \cap B \neq \emptyset\}.$$

Hence

$$\begin{aligned} \mu^*(A) + \mu^*(B) &\leq \sum_{i \in I} |Q_i| + \sum_{i \in J} |Q_i| \\ &\leq \sum_{i=1}^{\infty} |Q_i| < \mu^*(A \cup B) + \epsilon. \end{aligned}$$

Thus $\mu^*(A)+\mu^*(B) < \mu^*(A\cup B)+\epsilon$ for all $\epsilon > 0$ which implies that $\mu^*(A)+\mu^*(B) \leq \mu^*(A \cup B)$ and so $\mu^*(A) + \mu^*(B) = \mu^*(A \cup B)$.

iv) Fix a closed cuboid

$$Q = I_1 \times \cdots \times I_n, \ I_i = [a_i, b_i].$$

We claim that

$$|Q| \leq \mu^*(Q).$$

Equivalently, if $Q_i \in \mathcal{Q}_n, i \in \mathbb{N}$, is a sequence of closed cuboids, then

$$Q \subset \bigcup_{i=1}^{\infty} Q_i \Rightarrow |Q| \leq \sum_{i=1}^{\infty} |Q_i|. \tag{11.8}$$

For a closed interval $I = [a,b] \subset \mathbb{R}$ with $a < b$, we have

$$|I| = b - a.$$

Then

$$|I| - 1 \leq \sharp(I \cap \mathbb{Z}) \leq |I| + 1.$$

Hence

$$N|I| - 1 \leq \sharp(NI \cap \mathbb{Z}) \leq N|I| + 1.$$

and thus

$$|I| - \frac{1}{N} \leq \frac{1}{N}\sharp(I \cap \frac{1}{N}\mathbb{Z}) \leq |I| + \frac{1}{N}$$

for every integer $N \in \mathbb{N}$. Take the limit $N \to \infty$ to obtain

$$|I| = \lim_{N \to \infty} \frac{1}{N}\sharp(I \cap \frac{1}{N}\mathbb{Z}).$$

Thus

$$\begin{aligned} |Q| &= \lim_{N\to\infty} \prod_{i=1}^{n} \frac{1}{N} \sharp(I_i \cap \frac{1}{N}\mathbb{Z}) \\ &= \lim_{N\to\infty} \frac{1}{N^n} \sharp(Q \cap \frac{1}{N}\mathbb{Z}^n). \end{aligned} \tag{11.9}$$

Now suppose $Q_i \in \mathcal{Q}_n, i \in \mathbb{N}$, is a sequence of closed cuboids such that

$$Q \subset \bigcup_{i=1}^{\infty} Q_i.$$

Fix an $\epsilon > 0$ and choose a sequence of open cuboids $U_i \subset \mathbb{R}^n$ such that

$$Q_i \subset U_i, \;\; |U_i| < |Q_i| + \frac{\epsilon}{2^i}.$$

Since Q is compact and the U_i form an open cover of Q, there exists a $k \in \mathbb{N}$ such that

$$Q \subset \cup_{i=1}^{k} U_i.$$

This implies

$$\begin{aligned} \frac{1}{N^n}\sharp(Q \cap \frac{1}{N}\mathbb{Z}^n) &\leq \sum_{i=1}^{k} \frac{1}{N^n}\sharp(U_i \cap \frac{1}{N}\mathbb{Z}^n) \\ &\leq \sum_{i=1}^{k} \frac{1}{N^n}\sharp(\overline{U}_i \cap \frac{1}{N}\mathbb{Z}^n) \end{aligned}$$

Letting $N \to \infty$ and using (11.9), we get

$$|Q| \leq \sum_{i=1}^{k} |U_i| \leq \sum_{i=1}^{\infty} |U_i| \leq \sum_{i=1}^{\infty} \Big(|Q_i| + \frac{\epsilon}{2^i}\Big) = \epsilon + \sum_{i=1}^{\infty} |Q_i|. \tag{11.10}$$

Since $\epsilon > 0$ is arbitrarily, this proves the claim. Hence $\mu^*(Q) \leq |Q| \leq \mu^*(Q)$ and so $\mu^*(Q) = |Q|$. To prove that $\mu^*(Q^o) = |Q|$, fix a constant $\epsilon > 0$ and choose a closed cuboid $P \subset Q^o$ such that $|Q| < |P| + \epsilon$. Then

$$|Q| < |P| + \epsilon = \mu^*(P) + \epsilon \leq \mu^*(Q^o) + \epsilon.$$

Taking $\epsilon \to 0$, we have

$$|Q| \leq \mu^*(Q^o) \leq \mu^*(Q) = |Q|.$$

Hence $\mu^*(Q^o) = |Q|$. □

Definition 11.6 Let $\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0, \infty]$ be the Lebesgue outer measure. A subset $A \subset \mathbb{R}^n$ is called Lebesgue measurable if A is μ^*-measurable. The set of all Lebesgue measurable subsets of $\mathbb{R}^n$ will be denoted by

$$\mathcal{L} \equiv \mathcal{L}(\mathbb{R}^n) = \{A \subset \mathbb{R}^n : A \textit{ is Lebesgue measurable}\}.$$

The function

$$m := \mu^*|_{\mathcal{L}(\mathbb{R}^n)} : \mathcal{L}(\mathbb{R}^n) \to [0, \infty]$$

is called the Lebesgue measure on $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called Lebesgue measurable if it is measurable with respect to the Lebesgue σ-algebra $\mathcal{L}$ on $\mathbb{R}^n$ and the Borel σ-algebra on the target space $\mathbb{R}$.

Corollary 11.7 (i) $(\mathbb{R}^n, \mathcal{L}, m)$ is a complete measure space.

(ii) m is translation-invariant, i.e., if $A \in \mathcal{L}$ and $x \in \mathbb{R}^n$, then

$$m(A+x) = m(A).$$

(iii) Every Borel set in $\mathbb{R}^n$ is Lebesgue measurable.
(iv) If $Q \in \mathcal{Q}_n$, then

$$Q^o \in \mathcal{L}, \quad m(Q^o) = m(Q) = |Q|.$$

Theorem 11.8 (regularity of the Lebesgue outer measure). *The Lebesgue outer measure $\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0,\infty]$ satisfies the following.*
(i) For every subset $A \subset \mathbb{R}^n$,

$$\mu^*(A) = \inf\{\mu^*(U) : A \subset U \subset \mathbb{R}^n \textit{ and } U \textit{ open}\}.$$

(ii) If $A \subset \mathbb{R}^n$ is Lebesgue measurable, then

$$\mu^*(A) = \sup\{\mu^*(K) : K \subset A \textit{ and } K \textit{ is compact}\}.$$

(iii) If $\{A_i\}$ is an increasing sequence of subsets of $\mathbb{R}^n$, then

$$\mu^*(\cup_{i=1}^{\infty} A_i) = \lim_{i\to\infty} \mu^*(A_i).$$

Proof. (i) Fix a subset $A \subset \mathbb{R}^n$ and $\epsilon > 0$. The assertion is obvious when $\mu^*(A) = \infty$. Assume $\mu^*(A) < \infty$ and choose a sequence of closed cuboids $Q_i \in \mathcal{Q}_n$ such that

$$A \subset \bigcup_{i=1}^{\infty} Q_i, \ \sum_{i=1}^{\infty} |Q_i| \le \mu^*(A) + \frac{\epsilon}{2}.$$

Now choose a sequence of open cuboids $U_i \subset \mathbb{R}^n$ such that

$$Q_i \subset U_i, \ |U_i| \le |Q_i| + \frac{\epsilon}{2^{i+1}}.$$

Then $U := \bigcup_{i=1}^{\infty} U_i$ is an open subset of $\mathbb{R}^n$ containing A and

$$\mu^*(U) \le \sum_{i=1}^{\infty} |U_i| \le \sum_{i=1}^{\infty} \left(|Q_i| + \frac{\epsilon}{2^{i+1}}\right) \le \mu^*(A) + \epsilon.$$

This proves (i).

(ii) Assume first that $A \subset \mathbb{R}^n$ is Lebesgue measurable and bounded. Choose $r > 0$ so that $A \subset B_r := \{x \in \mathbb{R}^n : |x| < r\}$. Fix an $\epsilon > 0$. By (i), there exists an open set $U \subset \mathbb{R}^n$ such that

$$\overline{B_r} - A \subset U \textit{ and } \mu^*(U) \le \mu^*(\overline{B_r} - A) + \epsilon.$$

Hence $K := \overline{B_r} - U$ is a compact subset of A and K and U are disjoint measurable subsets. Thus we have

$$\begin{aligned} \mu^*(K) &= \mu^*(K \cup U) - \mu^*(U) \ge \mu^*(\overline{B_r}) - \mu^*(U) \\ &\ge \mu^*(\overline{B_r}) - (\mu^*(\overline{B_r} - A) + \epsilon) = \mu^*(A) - \epsilon. \end{aligned}$$

This proves (ii) for bounded Lebesgue measurable sets. Let $A \in \mathcal{L}(\mathbb{R}^n)$ be unbounded. Suppose that b is a real number less than $\mu^*(A)$; we will produce a compact subset K of A such that $b < \mu^*(K)$. Let $\{A_j\}$ be an increasing sequence of bounded measurable subsets of A such that $A = \cup_{j=1}^{\infty} A_j$. We have $\mu^*(A) = \lim_{j\to\infty} \mu^*(A_j)$, and so we can choose j_0 such that $\mu^*(A_{j_0}) > b$. Now apply part (ii) to A_{j_0} to get a compact subset K of A_{j_0} (and hence of A) such that $\mu^*(K) > b$. Since b was an arbitrary number less than $\mu^*(A)$, the proof is complete.

(iii) If $\mu^*(A_i) = \infty$ for some i, then the assertion is obvious. Hence assume $\mu^*(A_i) < \infty$ for all i and fix an $\epsilon > 0$. By (i), there is a sequence of open sets $U_i \subset \mathbb{R}^n$ such that $A_i \subset U_i$ and $\mu^*(U_i) \leq \mu^*(A_i) + \epsilon$ for all i. Taking $V_i = \cap_{k=i}^{\infty} U_k$, we have $V_i \uparrow$, $A_i \subset V_i$, $\mu^*(V_i) \leq \mu^*(U_i) \leq \mu^*(A_i) + \epsilon$. This yields

$$\mu^*(\cup_{i=1}^{\infty} A_i) \leq \mu^*(\cup_{i=1}^{\infty} V_i) = \lim_{i\to\infty} \mu^*(V_i) \leq \lim_{i\to\infty} \mu^*(A_i) + \epsilon$$

Since ϵ is arbitrary, the proof of (iii) is complete.

Theorem 11.9 *(the Lebesgue measure as a completion). Let $\mu^* : \mathcal{P}(\mathbb{R}^n) \to [0,\infty]$ be the Lebesgue outer measure, let $m = \mu^*|_{\mathcal{L}(\mathbb{R}^n)} : \mathcal{L}(\mathbb{R}^n) \to [0,\infty]$ be the Lebesgue measure, let $\mathcal{B} \subset \mathcal{L}(\mathbb{R}^n)$ be the Borel σ-algebra of $\mathbb{R}^n$, and define*

$$\nu = \mu^*|_{\mathcal{B}} : \mathcal{B} \to [0,\infty]$$

Then $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), m)$ is the completion of $(\mathbb{R}^n, \mathcal{B}, \nu)$.

Proof. Let $(\mathbb{R}^n, \mathcal{B}^*, \nu^*)$ denote the completion of $(\mathbb{R}^n, \mathcal{B}, \nu)$.

Claim. *Let $A \subset \mathbb{R}^n$. Then the following are equivalent.*
(I). $A \in \mathcal{A} := \mathcal{L}(\mathbb{R}^n)$.
(II). $A \in \mathcal{B}^*$, i.e., there exist Borel measurable sets $B_0, B_1 \in \mathcal{B}$ such that $B_0 \subset A \subset B_1$ and $\mu^*(B_1 - B_0) = 0$.

If the set A satisfies both (I) and (II), then

$$\mu^*(A) \leq \mu^*(B_1) = \mu^*(B_0) + \mu^*(B_1 - B_0) = \mu^*(B_0) \leq \mu^*(A);$$

and hence

$$m(A) = \mu^*(A) = \mu^*(B_0) = \nu^*(A).$$

This shows that $\mathcal{L}(\mathbb{R}^n) = \mathcal{B}^*$ and $m = \nu^*$. Thus it remains to prove the claim. Fix a subset $A \subset \mathbb{R}^n$. We show that (II) implies (I). Thus assume that $A \in \mathcal{B}^*$ and choose Borel measurable sets $B_0, B_1 \in \mathcal{B}$ such that

$$B_0 \subset A \subset B_1, \ \mu^*(B_1 - B_0) = 0.$$

Then $\mu^*(A - B_0) \leq \mu^*(B_1 - B_0) = 0$ and so $\mu^*(A - B_0) = 0$. Thus $A - B_0 \in \mathcal{A}$ and so $A = (A - B_0) \cup B_0 \in \mathcal{A}$.

We now show that (I) implies (II). Thus assume that $A \in \mathcal{A}$. Suppose first that $\mu^*(A) < \infty$. By Theorem 11.8, there exist a sequence of compact sets $K_i \subset \mathbb{R}^n$ and a sequence of open sets $U_i \subset \mathbb{R}^n$ such that

$$K_i \subset A \subset U_i, \ \mu^*(A) - \frac{1}{i} \leq \mu^*(K_i) \leq \mu^*(U_i) \leq \mu^*(A) + \frac{1}{i}.$$

Define

$$B_0 = \bigcup_{i=1}^{\infty} K_i, \ B_1 = \bigcap_{i=1}^{\infty} U_i.$$

These are Borel sets satisfying $B_0 \subset A \subset B_1$ and

$$\mu^*(A) - \frac{1}{i} \leq \mu^*(K_i) \leq \mu^*(B_0) \leq \mu^*(B_1) \leq \mu^*(U_i) \leq \mu^*(A) + \frac{1}{i}.$$

Taking the $i \to \infty$, we obtain

$$\mu^*(B_0) = \mu^*(B_1) = \mu^*(A) < \infty$$

and hence

$$\mu^*(B_1 - B_0) = 0.$$

This shows that $A \in \mathcal{B}^*$ for every $A \in \mathcal{A}$ with $\mu^*(A) < \infty$.

Now suppose that our set $A \in \mathcal{A}$ satisfies $\mu^*(A) = \infty$ and define

$$A_k := \{x \in A : |x| \le k\} \ for\ k \in \mathbb{N}.$$

From the above, $A_k \in \mathcal{B}^*$ for all k and so there exist sequences of Borel sets $\{B_k\}, \{B'_k\} \subset \mathcal{B}$ such that

$$B_k \subset A_k \subset B'_k,\ \mu^*(B'_k - B_k) = 0.$$

Define

$$B = \bigcup_{i=1}^{\infty} B_i,\ B' = \bigcup_{i=1}^{\infty} B'_i.$$

Then $B, B' \in \mathcal{B}, B \subset A \subset B'$, and

$$\mu^*(B' - B) \le \sum_{k=1}^{\infty} \mu^*(B'_k - B) \le \sum_{k=1}^{\infty} \mu^*(B'_k - B_k) = 0.$$

This shows that $A \in \mathcal{B}^*$ for every $A \in \mathcal{A}$. Thus we have proved that (I) implies (II) and this completes the proof of Theorem 11.9. □

Lemma 11.10 *If $f \in L^1(\mathbb{R}^n)$, there exists a sequence $\{h_k\}$ of continuous functions with compact support such that*

$$\int_{\mathbb{R}^n} |f - h_k| \to 0 \ as\ k \to \infty.$$

Proof. If f is an integrable function for which the above conclusion holds, we will say that f has property $\mathcal{P}$. We will prove the lemma by considering a series of special cases. We first observe that

(*) *a finite linear combination of functions with property $\mathcal{P}$ has property $\mathcal{P}$.*

We prove now that

(**) *if $\{f_k\}$ is a sequence of functions with property $\mathcal{P}$, and if $\int_{\mathbb{R}^n} |f_k - f| \to 0$, then f has property $\mathcal{P}$.*

Note that since f_k is integrable and $\int |f| \le \int |f - f_k| + \int |f_k|$, it follows that f is integrable. Next, given $\epsilon > 0$, choose k_0 so that $\int |f - f_{k_0}| < \frac{\epsilon}{2}$. Then choose a continuous h with compact support such that $\int |f_{k_0} - h| < \frac{\epsilon}{2}$. Since

$$\int |f - h| \le \int |f - f_{k_0}| + \int |f_{k_0} - h| < \epsilon,$$

we see that f has property $\mathcal{P}$.

To prove the lemma, let $f \in L^1(\mathbb{R}^n)$. Writing $f = f^+ - f^-$, we may assume by (*) that $f \ge 0$. Then, there exist nonnegative simple $s_k \uparrow f$. Thus, $s_k \in L^1(\mathbb{R}^n)$ and $\int |f - s_k| \to 0$, so that by (**), we may suppose that f is a nonnegative integrable simple function. Hence, by (*), we may assume that $f = \chi_E$ for a set E with $m(E) < \infty$. Given $\epsilon > 0$, choose an open G such that $E \subset G$ and $m(G - E) < \epsilon$. Then

$$\int |\chi_G - \chi_E| = m(G - E) < \epsilon.$$

so we may assume that $f = \chi_G$ for an open G with $m(G) < \infty$. We can write $G = \cup_{k=1}^{\infty} I_k$, where the I_k are disjoint, partly open cubes. If we let f_N be the characteristic function of $\cup_{k=1}^{N} I_k$, we obtain

$$\int |f - f_N| = \sum_{k=N+1}^{\infty} |I_k| \to 0,$$

since $\sum_{k=1}^{\infty} |I_k| = |G| < \infty$. By (**), it is thus enough to show that each f_N has property $\mathcal{P}$. But f_N is the sum of $\chi_{I_k}, k = 1, ..., N$, so it suffices by (*) to show that the characteristic function of any partly open cubes I has property $\mathcal{P}$. This is self-evident: if I^* denotes a closed cube that contains the closure of I in its interior and that satisfies $|I^* - I| < \epsilon$, then for any continuous h, $0 \le h \le 1$, which is 1 in I and 0 outside I^*, we have

$$\int |\chi_I - h| \le |I^* - I| < \epsilon.$$

This completes the proof of Lemma. $\square$

Theorem 11.11 *The set of continuous functions with compact support is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$.*

Proof. Suppose $f \in L^p$. We have $\int_{\mathbb{R}^n} |f - f\chi_{B(0,n)}|^p \to 0$ as $n \to \infty$ by the dominated convergence theorem, the dominating function being $|f|^p$. Hence it suffices to approximate functions in L^p that have compact support. By writing $f = f^+ - f^-$ we may suppose $f \ge 0$. Consider simple functions s_k increasing to f; then we have $\int_{\mathbb{R}^n} |f - s_k|^p \to 0$ by the dominated convergence theorem, so it suffices to approximate simple functions with compact support. By linearity, it suffices to approximate characteristic functions with compact support. By the same reasons as in the proof of Lemma 11.10, it suffices to approximate characteristic functions of any partly open cube. Also as in the proof of Lemma 11.10, given any partly open cube I, there exists g continuous with compact support and with values in $[0, 1]$ such that $\int_{\mathbb{R}^n} |g - \chi_I| < \epsilon$. Since $|g - \chi_I| \le 1$, then $\int_{\mathbb{R}^n} |g - \chi_I|^p \le \int_{\mathbb{R}^n} |g - \chi_I| < \epsilon$. This completes the proof. $\square$

The method of proof, where one proves a result for characteristic functions, then simple functions, then non-negative functions, and then finally integrable functions, is very common.

Theorem 11.12 (Continuity in L^p) *Let $f \in L^p(\mathbb{R}^n), 1 \le p < \infty$. Then*

$$\lim_{|h| \to 0} ||f(\cdot + h) - f(\cdot)||_p = 0.$$

Proof. Let C_p denote the class of $f \in L^p$ such that $||f(\cdot + h) - f(\cdot)||_p \to 0$ as $|h| \to 0$. We claim that (a) a finite linear combination of functions in C_p is in C_p, and (b) if $f_k \in C_p$ and $||f_k - f||_p \to 0$, then $f \in C_p$. Both of these facts follow easily from Minkowski's inequality; for (b), for example, note that

$$\begin{aligned} ||f(\cdot + h) - f(\cdot)||_p &\le ||f(\cdot + h) - f_k(\cdot + h)||_p + ||f_k(\cdot + h) - f_k(\cdot)||_p + ||f_k - f||_p \\ &= ||f_k(\cdot + h) - f_k(\cdot)||_p + 2||f_k - f||_p. \end{aligned}$$

Since $f_k \in C_p$, we have $\limsup_{|h|\to 0} ||f(\cdot+h)-f(\cdot)||_p \le 2||f_k - f||_p$, and (b) follows by letting $k \to \infty$. It is clear that the characteristic function of a cube belongs to C_p. From the proof of Theorem 10.11, we know that linear combinations of characteristic functions of cubes are dense in $L^p(\mathbb{R}^n)$, it follows from (a) and (b) that $L^p(\mathbb{R}^n)$ is contained in C_p, and the proof is complete.

□

Now we come to the transformation formula which describes how the integral of a Lebesgue measurable function transforms under composition with a C^1 diffeomorphism. Fix a positive integer $n \in \mathbb{N}$ and denote by $(\mathbb{R}^n, \mathcal{L}, m)$ the Lebesgue measure space. For any Lebesgue measurable set $X \subset \mathbb{R}^n$ denote by

$$\mathcal{L}_X = \{A \in \mathcal{L} : A \subset X\}$$

the restricted Lebesgue σ-algebra and by

$$m_X = m|_{\mathcal{L}_X} : \mathcal{L}_X \to [0, \infty]$$

the restriction of the Lebesgue measure to $\mathcal{L}_X$.

Theorem 11.13 (transformation formula). *Suppose $\phi : U \to V$ is a C^1 diffeomorphism between open subsets of $\mathbb{R}^n$.*

i) If $f : V \to [0, \infty]$ is Lebesgue measurable, then $f \circ \phi : U \to [0, \infty]$ is Lebesgue measurable and

$$\int_U (f \circ \phi)|det(d\phi)|dm = \int_V f dm.$$

ii) If $E \in \mathcal{L}_U$ and $f \in L^1(m_V)$, then $\phi(E) \in \mathcal{L}_V, (f \circ \phi)|det(d\phi)| \in L^1(m_U)$ and

$$\int_E (f \circ \phi)|det(d\phi)|dm = \int_{\phi(E)} f dm.$$

The proof of this theorem will be omitted.

Exercises 10

1. Show that $tB := \{tb : b \in B\}$ is a Borel set for all $B \in \mathcal{B}(\mathbb{R}^n)$ and $t > 0$.

2. Prove that for a compact subset A of $\mathbb{R}^n$, $m^*(A)$ can be computed using finite covers, i.e., $m^*(A) = \inf\{\sum_{k=1}^N |Q_i| : A \subset \cup_{i=1}^N Q_i\}$.

3. Show that for arbitrary sets $A \subset \mathbb{R}^n$, there exist $H \subset A \subset G$ such that H is F_σ, G is G_δ, and $|H| = \mu^*(A) = |G|$.

4. Let A, B be subsets of $\mathbb{R}^n$ with $\mu^*(A), \mu^*(B) < \infty$ such that $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. Prove: (a) $|A \cap B| = 0$. (b) If $A \cup B \in \mathcal{L}(\mathbb{R}^n)$, then $A, B \in \mathcal{L}(\mathbb{R}^n)$.

5. Let A, B be subsets of $\mathbb{R}^n$. Prove: (a) If $A \in \mathcal{L}(\mathbb{R}^n)$, then $\mu^*(A \cap B) + \mu^*(A \cup B) = |A| + \mu^*(B)$. (b) If there exists a measurable C such that $A \subset C$ and $|B \cap C| = 0$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

6. Let $A \subset \mathbb{R}^n$ with $\mu^*(A) < \infty$. Prove that $A \in \mathcal{L}(\mathbb{R}^n)$ iff given $\epsilon > 0$, there exists a finite collection of closed cubois $Q_1, ..., Q_N$, say, such that $\mu^*(A\Delta(\cup_{i=1}^N Q_i)) < \epsilon$.

7. Prove that if A is a compact subset of $\mathbb{R}^n$ and $O_k = \{x \in \mathbb{R}^n : d(x, A) < 1/k\}$, then $|A| = \lim_{k\to\infty} |O_k|$.

8. For $A \in \mathcal{L}(\mathbb{R}^n)$, let $f_A(x) = |A \cap B(0, |x|)|$. Find $\lim_{|x|\to\infty} f_A(x)$ and show that f_A is continuous on $\mathbb{R}^n$.

9. Let $A \in \mathcal{L}(\mathbb{R}^n)$ with $\mu^*(A) > 0$. Prove that for $r \in (0, \mu^*(A))$, there is a compact $K \subset A$ with $|K| = r$.

10. Let A be a measurable subset of $\mathbb{R}^n$ with positive finite measure. Prove that $\lim_{|x|\to 0} |A \cap (x + A)| = |A|$.

11. Let $A \subset \mathbb{R}^n$, $\mu^*(A) < \infty$, and $r \in \mathbb{R}$. Prove: (a) The dilation of A by r, $rA = \{rx : x \in A\}$, satisfies $\mu^*(rA) = |r|^n \mu^*(A)$. (b) A is measurable iff rA is measurable for some $r \neq 0$, and then $|rA| = |r|^n |A|$.

12. Let E be a given subset of $\mathbb{R}^n$. Show that the following five statements are equivalent:
(a) E is Lebesgue measurable;
(b) For every $\epsilon > 0$, there exists an open set $O \supset E$ such that $\mu^*(O - E) < \epsilon$;
(c) There exists a G_δ-set $G \supset E$ such that $\mu^*(G - E) = 0$;
(d) For every $\epsilon > 0$, there exists a closed set $F \subset E$ such that $\mu^*(E - F) < \epsilon$;
(e) There exists an F_σ-set $F \subset E$ such that $\mu^*(E - F) = 0$.

13. Let $A \subset E \subset \mathbb{R}^n$, where E is Lebesgue measurable and $|E| < \infty$. Show that A is measurable provided $|E| = \mu^*(E - A) + \mu^*(A)$.

14. If F is Lebesgue measurable and $\mu^*(F\Delta G) = 0$, show that G is Lebesgue measurable.

12. Product measures, Tonelli, Fubini theorems

In calculus one defines integrals over higher-dimensional regions and then evaluates these integrals by applying the usual techniques of integration, one variable at a time. Similar techniques work for the Lebesgue integral. That is, given σ-finite measures μ and ν on spaces X and Y, we can define a natural product measure on the product space $X \times Y$ Then we look at how integrals with respect to this product measure can be evaluated in terms of integrals with respect to μ and ν over X and Y

Definition 12.1 *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces, and, let $X \times Y$ be the Cartesian product of the sets X and Y. A subset of $X \times Y$ is a measurable rectangle if it has the form $A \times B$ for some $A \in \mathcal{A}$ and some $B \in \mathcal{B}$; the σ-algebra on $X \times Y$ generated by the collection of all measurable rectangles is called the product of the σ-algebras $\mathcal{A}$ and $\mathcal{B}$ and is denoted by $\mathcal{A} \otimes \mathcal{B}$.*

Example 12.2. Let us show that the product σ-algebra $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ is equal to $\mathcal{B}(\mathbb{R}^2)$. Recall that $\mathcal{B}(\mathbb{R}^2)$ is generated by the collection of all sets of the form

$(a, b] \times (c, d]$. Thus $\mathcal{B}(\mathbb{R}^2)$ is generated by a subfamily of the σ-algebra $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ and so is included in $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. We turn to the reverse inclusion. The projections π_1 and π_2 of $\mathbb{R}^2$ onto $\mathbb{R}$ defined by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$ are continuous and hence Borel measurable. It follows from this and the identity $A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B) = \pi_1^{-1}(A) \cap \pi_2^{-1}(B)$ that if A and B belong to $\mathcal{B}(\mathbb{R})$, then $A \times B$ belongs to $\mathcal{B}(\mathbb{R}^2)$. Since $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ is the σ-algebra generated by the collection of all such rectangles $A \times B$, it must be included in $\mathcal{B}(\mathbb{R}^2)$. Thus $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

Suppose that E is a subset of $X \times Y$. Then for each x in X and each y in Y the sections E_x and E_y are the subsets of Y and X given by

$$E_x = \{y \in Y : (x, y) \in E\}$$

and

$$E^y = \{x \in X : (x, y) \in E\}.$$

If f is a function on $X \times Y$, then the sections f_x and f^y are the functions on Y and X given by

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

and

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

Lemma 12.3. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces.*

(a) If E is a subset of $X \times Y$ that belongs to $\mathcal{A} \otimes \mathcal{B}$, then each section E_x belongs to $\mathcal{B}$ and each section E^y belongs to $\mathcal{A}$.

(b) If f is an extended real-valued (or a complex-valued) $A \otimes B$-measurable function on $X \times Y$, then each section f_x is $\mathcal{B}$-measurable and each section f^y is $\mathcal{A}$-measurable.

Proof. Let $\mathcal{F}$ be the collection of all subsets E of $X \times Y$ such that $E_x \in \mathcal{B}$. Then F contains all rectangles $A \times B$ for which $A \in \mathcal{A}$ and $B \in \mathcal{B}$ (note that $(A \times B)_x$ is either B or $\emptyset$). In particular, $X \times Y \in \mathcal{F}$. From $(E^c)_x = (E_x)^c$ and $(\cup_{n=1}^{\infty} E_n)_x = \cup_{n=1}^{\infty} (E_n)_x$, we know that $\mathcal{F}$ is closed under complementation and countable unions; thus $\mathcal{F}$ is a σ-algebra. It follows that $\mathcal{F}$ includes the σ-algebra $\mathcal{A} \otimes \mathcal{B}$ and hence that $E_x \in \mathcal{B}$ if $E \in \mathcal{A} \otimes \mathcal{B}$. A similar argument shows that $E_y \in \mathcal{A}$ if $E \in \mathcal{A} \otimes \mathcal{B}$. With this part (a) is proved. Part (b) follows from part (a) and the identities $(f_x)^{-1}(D) = (f^{-1}(D))_x$ and $(f^y)^{-1}(D) = (f^{-1}(D))^y$. □

Proposition 12.4 *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. If $E \in \mathcal{A} \otimes \mathcal{B}$, then the function $x \to \nu(E_x)$ is $\mathcal{A}$-measurable and the function $y \to \mu(E^y)$ is $\mathcal{B}$-measurable.*

Proof. First suppose that the measure ν is finite. Let $\mathcal{F}$ be the class of those sets $E \in \mathcal{A} \otimes \mathcal{B}$ for which the function $x \to \nu(E_x)$ is $\mathcal{A}$-measurable. (Lemma 12.3 implies that $E_x \in \mathcal{B}$, and hence that $\nu(E_x)$ is defined). If $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then $\nu((A \times B)_x) = \nu(B)\chi_A(x)$, and so the rectangle $A \times B \in \mathcal{F}$. In particular, $X \times Y \in \mathcal{F}$. Note that if $E, F \in \mathcal{A} \otimes \mathcal{B}$ such that $E \subset F$, then $\nu((F - E)_x) = \nu(F_x) - \nu(E_x)$, and that if $\{E_n\}$ is an increasing sequence of sets in $\mathcal{A} \otimes \mathcal{B}$, then $\nu((\cup_{n=1}^{\infty} E_n)_x) = \lim_{n \to \infty} \nu((E_n)_x)$; it follows that $\mathcal{F}$ is closed under proper differences and unions of increasing sequences of sets. Thus $\mathcal{F}$ is a d-system. Since the family of measurable rectangles is closed under finite intersections:

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2),$$

Theorem 4.5 implies that $\mathcal{F} = \mathcal{A} \otimes \mathcal{B}$. Thus $x \to \nu(E_x)$ is measurable for each $E \in \mathcal{A} \otimes \mathcal{B}$.

Now suppose that ν is σ-finite, and let $\{D_n\}$ be a sequence of disjoint subsets of Y that belong to B, have finite measure under ν, and satisfy $\cup_{n=1}^{\infty} D_n = Y$. Define finite measures $\nu_1, \nu_2, ...$ on $\mathcal{B}$ by letting $\nu_n(B) = \nu(B \cap D_n)$. Then for each n the function $x \to \nu_n(E_x)$ is $\mathcal{A}$-measurable; since $\nu(E_x) = \sum_{n=1}^{\infty} \nu_n(E_x)$ holds for each x, the measurability of $x \to \nu(E_x)$ follows. The function $y \to \mu(E^y)$ can be treated similarly, and so the proof is complete. □

Theorem 12.5. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces. Then there is a unique measure $\mu \times \nu$ on the σ-algebra $\mathcal{A} \otimes \mathcal{B}$ such that $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$ holds for each $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Furthermore, for any $E \in \mathcal{A} \otimes \mathcal{B}$, we have*

$$(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu = \int_Y \mu(E^y) d\nu. \tag{12.1}$$

The measure $\mu \times \nu$ is called the product of μ and ν.

Proof. We define functions $(\mu \times \nu)_1$, $(\mu \times \nu)_2$ on $\mathcal{A} \otimes \mathcal{B}$ by

$$(\mu \times \nu)_1(E) = \int_X \nu(E_x) d\mu \ , \ (\mu \times \nu)_2(E) = \int_Y \mu(E^y) d\nu$$

respectively. It is clear that $(\mu \times \nu)_1(\emptyset) = (\mu \times \nu)_2(\emptyset) = 0$. If $\{E_n\}$ is a sequence of disjoint sets in $\mathcal{A} \otimes \mathcal{B}$, $E = \cup_{n=1}^{\infty} E_n$, and if $x \in X$, then $\{(E_n)_x\}$ is a sequence of disjoint sets in $\mathcal{B}$ such that $E_x = \cup_{n=1}^{\infty} (E_n)_x$ and hence $\nu(E_x) = \sum_{n=1}^{\infty} \nu((E_n)_x)$. Thus

$$(\mu \times \nu)_1(E) = \int_X \nu(E_x) d\mu = \sum_{n=1}^{\infty} \int_X \nu((E_n)_x) d\mu = \sum_{n=1}^{\infty} (\mu \times \nu)_1(E_n),$$

and so $(\mu \times \nu)_1$ is countably additive. A similar argument shows that $(\mu \times \nu)_2$ is countably additive. Also if $A \in \mathcal{A}, \ B \in \mathcal{B}$, then

$$(\mu \times \nu)_1(A \times B) = \mu(A)\nu(B) = (\mu \times \nu)_2(A \times B).$$

Hence $(\mu \times \nu)_1$ and $(\mu \times \nu)_2$ are measures on $\mathcal{A} \otimes \mathcal{B}$ that have the same values on the measurable rectangles.

The uniqueness of $\mu \times \nu$ follows from Corollary 4.7. Thus $(\mu \times \nu)_1 = (\mu \times \nu)_2$, and Eq.(12.1) holds for each E in $\mathcal{A} \otimes \mathcal{B}$. □

The following two theorems enable one to evaluate integrals with respect to product measures by evaluating iterated integrals.

Theorem 12.6 (Tonelli's Theorem). *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $f : X \times Y \to [0, \infty]$ be $\mathcal{A} \otimes \mathcal{B}$-measurable. Then*

(a) the function $x \to \int_Y f_x d\nu$ is $\mathcal{A}$-measurable and the function $y \to \int_X f^y d\mu$ is $\mathcal{B}$-measurable, and

(b) f satisfies

$$\int_{X\times Y} fd(\mu\times\nu) = \int_X \left(\int_Y f_x d\nu\right) d\mu = \int_Y \left(\int_X f^y d\mu\right) d\nu. \tag{12.2}$$

Note that the functions f_x and f^y are nonnegative and measurable; thus the expression $\int_Y f_x d\nu$ is defined for each x in X and the expression $\int_X f^y d\mu$ is defined for each $y \in Y$.

Proof. First suppose that $E \in \mathcal{A}\otimes\mathcal{B}$ and that $f = \chi_E$. Then $f_x = \chi_{E_x}$ and $f^y = \chi_{E^y}$ and so the relations

$$\int_Y f_x d\nu = \nu(E_x) \textit{ and } \int_X f^y d\mu = \mu(E^y)$$

hold for each x and y. Thus Proposition 12.4 and Theorem 12.5 imply that conclusions (a) and (b) hold if f is a characteristic function. The additivity and homogeneity of the integral now imply that they hold for nonnegative simple $\mathcal{A}\otimes\mathcal{B}$-measurable functions. From Theorem 7.3, there is a sequence $\{f_n\}$ of simple $[0,\infty)$-valued $\mathcal{A}\otimes\mathcal{B}$-measurable functions that satisfy

$$f_1(x,y) \le f_2(x,y) \le \ldots \tag{12.3}$$

and

$$f(x,y) = \lim_{n\to\infty} f_n(x,y) \tag{12.4}$$

at each (x,y) in $X\times Y$. It then follows from the monotone convergence theorem (Theorem 8.4) that conclusions (a) and (b) hold for nonnegative $\mathcal{A}\otimes\mathcal{B}$-measurable functions. □

Example 12.7 Tonelli's Theorem can fail without the hypothesis of σ-finiteness.
Suppose $\mathcal{B}$ is the σ-algebra of Borel subsets of $[0,1]$, m is Lebesgue measure on $([0,1],\mathcal{B})$, and μ is counting measure on $([0,1],\mathcal{B})$. Let $D = \{(x,x) : x \in [0,1]\}$; then

$$\int_{[0,1]}\int_{[0,1]} \chi_D d\mu dm = \int_{[0,1]} dm = 1$$

and

$$\int_{[0,1]}\int_{[0,1]} \chi_D dm d\mu = \int_{[0,1]} d\mu = 0.$$

Theorem 12.8 (Fubini's Theorem). *Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be σ-finite measure spaces, and let $f : X\times Y \to [-\infty,\infty]$ be $\mathcal{A}\otimes\mathcal{B}$-measurable and*

$$\int_{X\times Y} |f| d(\mu\times\nu) < \infty.$$

Then

$$\int_Y |f(x,y)| d\nu < \infty \textit{ for a.e. } x \in X,$$

and

$$\int_X |f(x,y)| d\mu < \infty \textit{ for a.e. } y \in Y.$$

Furthermore,

$$a) \quad x \to \int_Y f(x,y)d\nu \text{ is an } \mathcal{A}-\text{measurable function on } X,$$

$$b) \quad y \to \int_X f(x,y)d\mu \text{ is a } \mathcal{B}-\text{measurable function on } Y$$

and

$$\int_{X\times Y} fd(\mu\times\nu) = \int_X \left(\int_Y f_x d\nu\right) d\mu = \int_Y \left(\int_X f^y d\mu\right) d\nu. \tag{12.5}$$

Proof. Let f^+ and f^- be the positive and negative parts of f; then Lemma 12.3 implies that $f_x, (f^+)_x$, and $(f^-)_x$ are $\mathcal{B}$-measurable, and Tonelli's theorem implies that

$$\int_X \left(\int_Y |f(x,y)|d\nu\right) d\mu < \infty$$

and so $\int_Y |f(x,y)|d\nu < \infty$ for a.e. x. We also know from Tonelli theorem that $x \to \int_Y (f^+)_x d\nu$ and $x \to \int_Y (f^-)_x d\nu$ are $\mathcal{A}$-measurable and are a.e. finite functions on X. Thus $x \to \int_Y f_x d\nu$ is $\mathcal{A}$-measurable and $\int_Y f_x d\nu = \int_Y (f^+)_x d\nu - \int_Y (f^-)_x d\nu$. From Tonelli theorem, we have

$$\int_{X\times Y} f^+ d(\mu\times\nu) = \int_X \left(\int_Y (f^+)_x d\nu\right) d\mu = \int_Y \left(\int_X (f^+)^y d\mu\right) d\nu. \tag{12.6}$$

The rest is obvious. □

Example 12.9 Let $(X, \mathcal{A}, \mu)$ be a σ -finite measure space, let m be Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and let $f : X \to [0,\infty]$ be $\mathcal{A}$-measurable. Let E be defined by

$$E = \{(x,t) \in X\times\mathbb{R} : 0 \le t < f(x)\}; \tag{12.7}$$

E is the "region under the graph of f". Then $E \in \mathcal{A}\otimes\mathcal{B}(\mathbb{R})$. In fact, if $f = \sum_{i=1}^n c_i\chi_{A_i}$ is a simple measurable function, then

$$E = \left\{(x,t) \in X\times[0,\infty] : t < \sum_{i=1}^n c_i\chi_{A_i}(x)\right\} = \cup_{i=1}^n (A_i\times[0,c_i)) \in \mathcal{A}\otimes\mathcal{B}(\mathbb{R}).$$

When $f : X \to [0,\infty]$ is any $\mathcal{A}$-measurable function, there exists a sequence of simple measurable functions $s_n : X \to [0,\infty)$ such that for every x in X, $0 \le s_1(x) \le s_2(x) \le \cdots$; and $\lim_{n\to\infty} s_n(x) = f(x)$. From the above, each $\{(x,t) \in X\times[0,\infty] : t < s_n(x)\}$ is in $\mathcal{A}\otimes\mathcal{B}(\mathbb{R})$, so

$$\{(x,t) \in X\times\mathbb{R} : 0 \le t < f(x)\} = \cup_{n=1}^\infty \{(x,t) \in X\times[0,\infty] : t < s_n(x)\}$$

is $\mathcal{A}\otimes\mathcal{B}(\mathbb{R})$-measurable. We can compute $\mu\times m(E)$ by using Tonelli theorem. On the one hand,

$$\mu\times m(E) = \int_X m(E_x)d\mu = \int_X fd\mu.$$

On the other hand,

$$\mu\times m(E) = \int_{\mathbb{R}} \mu(E^t)dt = \int_0^\infty \mu(\{x\in X : f(x) > t\})dt.$$

Thus we have the often useful relation

$$\int_X f d\mu = \int_0^\infty \mu(\{x \in X : f(x) > t\})dt. \tag{12.8}$$

The next result is a general theorem about boundedness of integral operators on L^p spaces.

Theorem 12.10 *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let K be a $\mathcal{A} \otimes \mathcal{B}$-measurable function on $X \times Y$. Suppose that there exists $C > 0$ such that $\int_X |K(x,y)| d\mu \le C$ for a.e. $y \in Y$ and $\int_X |K(x,y)| d\nu \le C$ for a.e. $x \in X$, and that $1 < p < \infty$. If $f \in L^p(\nu)$, the integral*

$$(Tf)(x) = \int_Y K(x,y) f(y) d\nu$$

converges absolutely for a.e. $x \in X$, the function Tf thus defined is in $L^p(\mu)$, and $||Tf||_p \le C||f||_p$.

Proof. Let q be the conjugate exponent to p. By applying Hölder's inequality to the product

$$|K(x,y)f(y)| = |K(x,y)|^{1/q}(|K(x,y)|^{1/p}|f(y)|)$$

We have

$$\begin{aligned} \int_Y |K(x,y)f(y)| d\nu &\le \left(\int_Y |K(x,y)| d\nu\right)^{1/q} \left(\int_Y |K(x,y)||f(y)|^p d\nu\right)^{1/p} \\ &\le C^{1/q} \left(\int_Y |K(x,y)||f(y)|^p d\nu\right)^{1/p} \end{aligned}$$

for a. e. $x \in X$. Hence, by Tonelli's theorem,

$$\begin{aligned} \int_X \left(\int_Y |K(x,y)f(y)| d\nu\right)^p d\mu &\le C^{p/q} \int_X \int_Y |K(x,y)||f(y)|^p d\nu d\mu \\ &\le C^{(p/q)+1} \int_Y |f(y)|^p d\nu. \end{aligned}$$

Since the last integral is finite, we know that $K(x, \cdot)f \in L^1(\nu)$ for a.e. x, so that Tf is well defined a.e., and

$$\int_X |(Tf)(x)|^p d\mu \le C^{(p/q)+1} ||f||_p^p.$$

Taking pth roots, we are done.

□

Minkowski's inequality states that the L^p norm of a sum is at most the sum of the L^p norms. There is a generalization of this result in which sums are replaced by integrals:

Theorem 12.11 *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let f be a nonnegative $\mathcal{A} \otimes \mathcal{B}$-measurable function on $X \times Y$. If $1 \le p < \infty$, then*

$$\left[\int_X \left(\int_Y f(x,y) d\nu\right)^p d\mu\right]^{1/p} \le \int_Y \left(\int_X f(x,y)^p d\mu\right)^{1/p} d\nu. \tag{12.9}$$

Proof. If $p = 1$, (12.9) follows from Tonelli's theorem. If $1 < p < \infty$, let q be the conjugate exponent to p and suppose $g \in L^q(\mu)$. The inequality (12.9) is certainly true if the right hand side is ∞; so without loss of generality, assume right hand side of (12.9) is finite.

Then by Tonelli's theorem and Hölder's inequality,

$$\begin{aligned} \int_X \left(\int_Y f(x,y) d\nu \right) |g(x)| d\mu &= \int_Y \left(\int_X f(x,y) |g(x)| d\mu \right) d\nu \\ &\leq ||g||_q \int_Y \left(\int_X f(x,y)^p d\mu \right)^{1/p} d\nu < \infty. \end{aligned}$$

So the map $T : L^q(\mu) \to \mathbb{R}$ defined by

$$T(g) = \int_X \left(\int_Y f(x,y) d\nu \right) g(x) d\mu$$

is a well-defined (i.e., the integral over with respect to $d\mu$ is defined), bounded linear functional on $L^q(\mu)$, and

$$||T|| \leq \int_Y \left(\int_X f(x,y)^p d\mu \right)^{1/p} d\nu.$$

On the other hand, it follows from the Riesz representation theorem that

$$||T|| = \left(\int_X \left(\int_Y f(x,y) d\nu \right)^p d\mu \right)^{1/p}.$$

Thus (12.9) holds. □

Exercises 11

1. Let $\mathcal{M}$ be an algebra of subsets of X, $\mathcal{N}$ an algebra of subsets of Y, and $\mathcal{R} = \{A \times B \subset X \times Y : A \in \mathcal{M}, B \in \mathcal{N}\}$ the rectangles in $\mathcal{M} \times \mathcal{N}$. Prove that the collection of finite unions of rectangles in $\mathcal{M}$ form an algebra of subsets of $A \times Y$.

2. Show that $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m) = \mathcal{B}(\mathbb{R}^{n+m})$.

3. Show that $\mathcal{B}(\mathbb{R}^{n+m}) \subset \mathcal{L}(\mathbb{R}^n) \otimes \mathcal{L}(\mathbb{R}^m)$ and that $\mathcal{L}(\mathbb{R}^{n+m})$ is the completion of $\mathcal{L}(\mathbb{R}^n) \otimes \mathcal{L}(\mathbb{R}^m)$.

4. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be probability measure spaces and $E \in \mathcal{A} \otimes \mathcal{B}$ such that $\mu \otimes \nu(E) = a^2 < 1$. Let $A = \{x \in X : \mu(E_x) \geq a\}$. Prove that $\mu(A) \leq a$.

5. Let $X = Y = [0,1]$, equipped with the σ-algebra of Lebesgue measurable subsets and Lebesgue measure. Consider a real sequence $\{a_n\}$, $0 < a_1 < a_2 < \cdots$, satisfying $\lim_{n\to\infty} a_n = 1$. For each n, choose a continuous function g_n such that $\{t : g_n(t) \neq 0\} \subset (a_n, a_{n+1})$ and also $\int_{[0,1]} g_n(t) dt = 1$. Define

$$f(x,y) = \sum_{n=1}^{\infty} (g_n(x) - g_{n+1}(x)) g_n(y).$$

Note that for each (x, y), at most two terms in the sum can be nonzero. Thus no convergence problem arises in the definition of f. Show that f is not integrable and that its repeated integrals do not agree.(Hints Measure 583)

6. Let $(X, \mathcal{A}, \mu)$ be a measure space and for $A \subset X$ and $0 < a < \infty$, put $A_a = \{(x, y) : a \in A, 0 < y < a\}$ for finite a and $A_\infty = \{(x, y) : x \in A, 0 < y < \infty\}$. Prove that if $A \in \mathcal{A}$, then $A_a \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R})$ and compute $\mu \times m(A_a)$.

7. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and $f \in L^1(\mu), g \in L^1(\nu)$. Prove that $h(x, y) = f(x)g(y) \in L^1(\mu\times\nu)$ and $\int_{X\times Y} hd(\mu\times\nu) = (\int_X fd\mu)(\int_Y fd\nu)$.

8. Let $(X, \mathcal{A}, \mu)$ be a σ-finite measure space and $f \in L^1$. Show that $\int_X fd\mu = \int_0^\infty \mu(\{f > t\})dt - \int_0^\infty \mu(\{f < t\})dt$.

13. The maximal theorem

In this section we shall define the Hardy-Littlewood maximal function and prove the Maximal theorem. We recall the following

Lemma 13.1 *Suppose E is a Lebesgue measurable set in $\mathbb{R}^n$. Then*

$$m(E) = \inf\{m(U) : E \subset U, U \text{ open}\} = \sup\{m(K) : K \subset E, K \text{ compact}\}.$$

Lemma 13.2 *Let $\mathcal{C}$ be a collection of open balls in $\mathbb{R}^n$ and let $U = \cup_{B\in\mathcal{C}} B$. If $m(U) > c$, then there exist disjoint $B_1, \cdots, B_k \in \mathcal{C}$, such that*

$$\sum_{i=1}^k m(B_i) > \frac{c}{3^n}.$$

Proof. If $m(U) > c$, by Lemma 13.1 there is a compact $K \subset U$ with $m(K) > c$. Since $\{B \in \mathcal{C}\}$ is an open cover of K, we can find finitely many of the balls in $\mathcal{C}$, say, $A_1, ..., A_m$, which cover K. Let B_1 be the largest of the A_j's (that is, choose B_1 to have maximal radius), let B_2 be the largest of the A_j's that are disjoint from B_1, B_3 the largest of the A_j's that are disjoint from B_1, and B_2, and so on until the list of A_j's is exhausted. According to this construction, if A_i is not one of the B_j's, there is a j such that $A_i \cap B_j \neq \emptyset$, and if j is the smallest integer with this property, the radius of A_i is at most that of B_j. Hence $A_i \subset B_j^*$, where B_j^* is the ball concentric with B_j whose radius is three times that of B_j. But then $K \subset \cup_{j=1}^m A_j \subset \cup_{i=1}^k B_i^*$, so

$$c < m(K) \leq \sum_{i=1}^k m(B_i^*) = 3^n \sum_{i=1}^k m(B_i).$$

□

Definition 13.3 A measurable function $f : \mathbb{R}^n \to \mathbb{C}$ is called locally integrable (with respect to Lebesgue measure) if $\int_K |f|dm < \infty$ for every bounded measurable set $K \subset \mathbb{R}^n$.

We denote the space of locally integrable functions by L^1_{loc}. If $f \in L^1_{loc}$, $x \in \mathbb{R}^n$, and $r > 0$, we define $A_r f(x)$ to be the average value of f on $B(x, r)$:

$$A_r f(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} f dm.$$

Definition 13.4 With each $f \in L^1_{loc}(\mathbb{R}^n, m)$, we associate its maximal function, f^*, which is defined as

$$f^*(x) = \sup_{r>0} A_r|f|(x) := \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f| dm.$$

Note that from $f \in L^1_{loc}$, for each fixed $r > 0$,

$$x \to \frac{1}{m(B(x,r))} \int_{B(x,r)} |f| dm$$

is a continuous function of x, which implies that $\{f^* > t\}$ is an open set for each real number t. Thus f^* is Lebesgue measurable.

Theorem 13.5 (Hardy-Littlewood). *There is a constant $C > 0$ such that for all $f \in L^1(\mathbb{R}^n, m)$ and all $\alpha > 0$,*

$$m(\{x \in \mathbb{R}^n : f^*(x) > \alpha\}) \leq \frac{C}{\alpha} \int_{\mathbb{R}^n} |f| dm. \tag{13.1}$$

Proof. Let $E_\alpha = \{x \in \mathbb{R}^n : \ f^*(x) > \alpha\}$. For each $x \in E_\alpha$, we can choose $r_x > 0$ such that $A_{r_x}|f|(x) > \alpha$. The balls $B(x, r_x)$ cover E_α and so by Lemma 13.2, if $c < m(E_\alpha)$ there exist $x_1, ..., x_k \in E_\alpha$ such that the balls $B_j = B(x_j, r_{x_j})$ are disjoint and $\sum_{j=1}^k m(B_j) > 3^{-n} c$. But then

$$c < 3^n \sum_{j=1}^k m(B_j) \leq \frac{3^n}{\alpha} \sum_{j=1}^k \int_{B_j} |f| dm \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f| dm.$$

Letting $c \to m(E_\alpha)$, we obtain the desired result. □

We shall use the notion of limit superior for real-valued functions of a real variable,

$$\limsup_{r \to t} f(r) = \lim_{\epsilon \to 0} \sup_{0<|r-t|<\epsilon} f(r) = \inf_{\epsilon > 0} \sup_{0<|r-t|<\epsilon} f(r).$$

and the easily verified fact that

$$\lim_{r \to t} f(r) = c \Leftrightarrow \limsup_{r \to t} |f(r) - c| = 0.$$

Theorem 13.6 *If $f \in L^1_{loc}$, then $\lim_{r \to 0} A_r f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$.*

Proof. It suffices to show that for $N \in \mathbb{N}$, $\lim_{r \to 0} A_r f(x) = f(x)$ for a.e. x with $|x| < N$. But for $|x| < N$ and $r < 1$ the values $A_r f(x)$ depend only on the values $f(y)$ for $|y| < N + 1$, so by replacing f with $f\chi_{B(0,N+1)}$, we may assume that $f \in L^1$.

Given $\epsilon > 0$, we can find a continuous integrable function g such that

$$\int_{\mathbb{R}^n} |g(y) - f(y)|dy < \epsilon.$$

Continuity of g implies that for every fixed $x \in \mathbb{R}^n$ and $\tilde{\epsilon} > 0$ there exists $\delta > 0$ such that $|y - x| < \delta \Rightarrow |g(y) - g(x)| < \tilde{\epsilon}$ and so for $r \in (0, \delta)$ we have

$$\begin{aligned} |A_r g(x) - g(x)| &= \frac{1}{m(B(x,r))} \left| \int_{B(x,r)} (g(y) - g(x))dy \right| \\ &\leq \frac{1}{m(B(x,r))} \int_{B(x,r)} |g(y) - g(x)|dy < \tilde{\epsilon}. \end{aligned}$$

Therefore $A_r g(x) \to g(x)$ as $r \to 0$ for every x, which implies that

$$\begin{aligned} & \limsup_{r\to 0} |A_r f(x) - f(x)| \\ = & \limsup_{r\to 0} |A_r(f-g)(x) + (A_r g - g)(x) + (g-f)(x)| \\ \leq & (f-g)^*(x) + 0 + |(f-g)|(x) = (f-g)^*(x) + |(f-g)|(x). \end{aligned}$$

Hence, if

$$E_\alpha = \{x : \limsup_{r\to 0} |A_r f(x) - f(x)| > \alpha\}, \ F_\alpha = \{x : |f-g|(x) > \alpha\},$$

we have

$$E_\alpha \subset F_{\frac{\alpha}{2}} \cup \{x : (f-g)^*(x) > \frac{\alpha}{2}\}.$$

But

$$\frac{\alpha}{2} m(F_{\frac{\alpha}{2}}) \leq \int_{F_{\frac{\alpha}{2}}} |f(x) - g(x)|dx < \epsilon,$$

so by the maximal theorem,

$$m(E_\alpha) \leq \frac{2\epsilon}{\alpha} + \frac{2C\epsilon}{\alpha}.$$

Since $\epsilon > 0$ is arbitrary, $m(E_\alpha) = 0$ for all $\alpha > 0$. But $\lim_{r\to 0} A_r f(x) = f(x)$ for all $x \notin \cup_{n=1}^\infty E_{\frac{1}{n}}$, we are done. $\square$

This result can be rephrased as follows: If $f \in L^1_{loc}$, then

$$\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} (f(y) - f(x))dy = 0 \ for \ a.e.x. \tag{13.2}$$

Something stronger is true: the above inequality remains valid if one replaces the integrand by its absolute value. That is, let us define the Lebesgue set L_f of f to be

$$L_f = \left\{ x : \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy = 0 \right\}.$$

Theorem 13.7 *If $f \in L^1_{loc}$, then $m((L_f)^c) = 0$.*

Proof. For each $c \in \mathbb{C}$ we can apply Theorem 13.6 to $g_c(x) = |f(x) - c|$ to conclude that, except on a Lebesgue null set E_c,

$$\lim_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c| dy = |f(x) - c|.$$

Let D be a countable dense subset of $\mathbb{C}$, and let $E = \cup_{c \in D} E_c$. Then $m(E) = 0$, and if $x \notin E$, for any $\epsilon > 0$ we can choose $c \in D$ such that $|f(x) - c| < \epsilon$, so that $|f(y) - f(x)| < |f(y) - c| + \epsilon$. Thus

$$\limsup_{r \to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| dy \leq |f(x) - c| + \epsilon < 2\epsilon.$$

Since ϵ is arbitrary, the desired result follows. □

Theorem 13.8 *Let $1 < p \leq \infty$ and $f \in L^p(\mathbb{R}^n)$. Then $f^* \in L^p(\mathbb{R}^n)$ and $||f^*||_p \leq c||f||_p$, where c depends only on n and p.*

Proof. Let $f \in L^p(\mathbb{R}^n)$. We may assume that $1 < p < \infty$ since the result is obvious with constant $c = 1$ when $p = \infty$. The idea is to obtain information for L^p by interpolating between the known results for L^1 and L^∞ . For $\alpha > 0$, let

$$h(\alpha) = |\{x \in \mathbb{R}^n : f^*(x) > \alpha\}|$$

denote the distribution function of f^*. Fix $\alpha > 0$ and define a function g by $g(x) = f(x)$ when $|f(x)| \geq \alpha/2$ and $g(x) = 0$ otherwise. Note that $g \in L^1(\mathbb{R}^n)$ since

$$\begin{aligned} ||g||_1 &= \int_{\{x \in \mathbb{R}^n : |f(x)| \geq \alpha/2\}} |f(x)| dx \\ &\leq \int_{\mathbb{R}^n} |f(x)| \left(\frac{|f(x)|}{\alpha/2} \right)^{p-1} dx = \left(\frac{2}{\alpha} \right)^{p-1} ||f||_p < \infty. \end{aligned}$$

Also, we have $||f - g||_\infty \leq \alpha/2$. Then since $|f(x)| \leq |g(x)| + \alpha/2$,

$$f^*(x) \leq g^*(x) + \frac{\alpha}{2}.$$

In particular,

$$\{x \in \mathbb{R}^n : f^*(x) > \alpha\} \subset \{x \in \mathbb{R}^n : g^*(x) > \alpha/2\},$$

so that, by (13.1),

$$\begin{aligned} h(\alpha) &\leq |\{x \in \mathbb{R}^n : g^*(x) > \alpha/2\}| \\ &\leq \frac{2C}{\alpha} ||g||_1 = \frac{2C}{\alpha} \int_{\{x \in \mathbb{R}^n : |f(x)| \geq \alpha/2\}} |f(x)| dx. \end{aligned} \tag{13.3}$$

We have

$$\int_{\mathbb{R}^n} f^{*p} dx = p \int_0^\infty \alpha^{p-1} h(\alpha) d\alpha.$$

Hence,

$$\int_{\mathbb{R}^n} f^{*p} dx \leq p \int_0^\infty \alpha^{p-1} \left[\frac{2C}{\alpha} \int_{\{x \in \mathbb{R}^n : |f(x)| \geq \alpha/2\}} |f(x)| dx \right] d\alpha.$$

Interchanging the order of integration in the expression on the right (which is justified since the integrand is nonnegative), we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} f^{*p}dx &\leq 2Cp\int_{\mathbb{R}^n}|f(x)|\left(\int_0^{2|f(x)|}\alpha^{p-2}d\alpha\right)dx \qquad (13.4)\\ &= \frac{2^pCp}{p-1}\int_{\mathbb{R}^n}|f(x)|^pdx = \frac{2^pCp}{p-1}||f||_p^p.\end{aligned}$$

Taking pth roots, we see that $||f^*||_p \leq c||f||_p$. □

Exercises 12

1. Show that if $f \in L^1(\mathbb{R}^n), f \neq 0$, then there exist $C,\ R > 0$ such that $f^*(x) > C|x|^{-n}$ for $|x| > R$.

2. Define

$$H^*f(x) = \sup\left\{\frac{1}{|B|}\int_B |f(y)|dy :\ B \textit{ is a ball and } x \in B\right\}.$$

Show that $f^* \leq H^*f \leq 2^nf^*$.

3. Show that if $f \in L^1_{\mathrm{loc}}$ and f is continuous at x, then x is in the Lebesgue set of f.

4. If E is a Borel set in $\mathbb{R}^n$, the density $D_E(x)$ of E at x is defined as

$$D_E(x) = \lim_{r\to 0}\frac{|E\cap B(x,r)|}{|B(x,r)|},$$

whenever the limit exists. Show that $D_E(x) = 1$ for a.e. $x \in E$ and $D_E(x) = 0$ for a.e. $x \in E^c$.

14. Riese-Thorin and Marcinkiewicz interpolation theorems

If $1 < p < q < r \leq \infty$, then $L^p \cap L^r \subset L^q \subset L^p + L^r$, and it is natural to ask whether a linear operator T on $L^p + L^r$ that is bounded on both L^p and L^r is also bounded on L^q. The answer is affirmative. The two fundamental theorems on this question are the Riesz-Thorin and Marcinkiewicz interpolation theorems, which we present in this section. We begin with the Riesz-Thorin theorem, whose proof is based on the following result from complex function theory.

Lemma 14.1 (The Three Lines Lemma) *Let h be a bounded continuous function on the strip $0 \leq Re\ z \leq 1$ that is holomorphic on the interior of the strip. If $|h| \leq M_0$ for $Re\ z = 0$ and $|h(z)| \leq M_1$ for $Re\ z = 1$, then $|h| \leq M_0^{1-t}M_1^t$ for $Re\ z = t, 0 < t < 1$.*

Proof. We can assume that $M_0M_1 \neq 0$. For $\epsilon > 0$, let

$$h_\epsilon(z) = h(z)M_0^{z-1}M_1^{-z}e^{\epsilon z(z-1)}.$$

Then h_ϵ is holomorphic inside the strip. We have

$$|h_\epsilon(iy)| \leq |M_0^{iy}M_1^{-iy}e^{\epsilon iy(iy-1)}| \leq e^{-\epsilon y^2} \leq 1.$$

Similarly,

$$|h_\epsilon(1+iy)| \leq |M_0^{1+iy-1} M_1^{-iy} e^{\epsilon(1+iy)(1+iy-1)}| \leq 1.$$

Now for any $x \in (0,1)$,

$$\begin{aligned} |h_\epsilon(x+iy)| &= |h(x+iy) M_0^{x+iy-1} M_1^{-x-iy} e^{\epsilon(x+iy)(x+iy-1)}| \\ &= |h(x+iy) M_0^{x-1} M_1^{-x} e^{\epsilon x(x-1)} e^{-\epsilon y^2}| \\ &\leq Ce^{-\epsilon y^2} \end{aligned}$$

for some positive constant C. This follows from the fact that h and $e^{\epsilon x(x-1)}$ are bounded on the strip. Since $e^{-\epsilon y^2} \to 0$ as $|y| \to \infty$, it follows that $|h_\epsilon(x+iy)| \to 0$ as $|y| \to \infty$. Thus, we can pick A large enough so that $|h_\epsilon(z)| \leq 1$ on the boundary of the rectangle $0 \leq \text{Re } z \leq 1$ and $-A \leq \text{ Im } z \leq A$. Then the maximum modulus principle implies that $|h_\epsilon(z)| \leq 1$ on the interior of the rectangle as well. Since this holds for any arbitrarily large A, we have that $|h_\epsilon(z)| \leq 1$ on the entire strip. Letting $\epsilon \to 0$, we obtain

$$|h(z)| M_0^{t-1} M_1^{-t} = \lim_{\epsilon \to 0} |h_\epsilon(z)| \leq 1 \; for \; Re \; z = t.$$

□

Definition 14.2 A measure μ on a measurable space $(X, \mathcal{A})$ is semifinite if for any measurable A with $\mu(A) > 0$, there exists a measurable $B \subset A$ such that $0 < \mu(B) < \infty$.

Theorem 14.3 *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $1 \leq p,\ q \leq \infty$ be conjugate exponents. Suppose that g is a measurable function on X such that $fg \in L^1$ for all f in the space $\mathcal{S}$ of simple measurable functions that vanish outside a set of finite measure, and the quantity*

$$\mathcal{M}_q(g) = \sup \left\{ \left| \int fg \right|,\ f \in \mathcal{S} \text{ and } ||f||_p = 1 \right\}$$

is finite. Also, suppose that $S_g = \{x : g(x) \neq 0\}$ is σ-finite if $q < \infty$ and that μ is semifinite if $q = \infty$. Then $g \in L_q$ and $\mathcal{M}_q(g) = ||g||_q$.

Proof. First, we remark that if f is a bounded measurable function that vanishes outside a set E of finite measure and $||f||_p = 1$, then $|\int fg| \leq \mathcal{M}_q(g)$. Indeed, there is a sequence $\{f_n\}$ of simple measurable functions such that $|f_n| \leq |f|$ (in particular, f_n vanishes outside E) and $f_n \to f$ pointwise. Since $|f_n g| \leq ||f||_\infty \chi_E |g|$ and $\chi_E |g| \in L^1$, by the dominated convergence theorem we have

$$\begin{aligned} \left| \int fg \right| &= \lim_{n \to \infty} \left| \int f_n g \right| \\ &\leq \lim_{n \to \infty} ||f_n||_p \mathcal{M}_q(g) = ||f||_p \mathcal{M}_q(g) = \mathcal{M}_q(g). \end{aligned}$$

Suppose that $q < \infty$. Let $\{E_n\}$ be an increasing sequence of sets of finite measure such that $S_g = \cup_{n=1}^\infty E_n$. Let $\{s_n\}$ be a sequence of simple measurable functions such that $s_n \to g$ pointwise and $|s_n| \leq |g|$ and let $g_n = s_n \chi_{E_n}$. Then $g_n \to g$ pointwise, $|g_n| \leq |g|$, and g_n vanishes outside E_n. Let

$$f_n = \frac{|g_n|^{q-1} \overline{\text{sng } g}}{||g_n||_q^{q-1}},$$

where

$$\operatorname{sng}(g) = \begin{cases} \frac{g}{|g|}, & if\ g \neq 0, \\ 0, & if\ g = 0. \end{cases}$$

Note that $g\ \overline{\operatorname{sng}(g)} = |g|$.

We have $||f_n||_p = 1$, and by Fatou's lemma,

$$\begin{aligned} ||g||_q &\leq \liminf_{n\to\infty} ||g_n||_q = \liminf_{n\to\infty} \int |f_n g_n| \\ &\leq \liminf_{n\to\infty} \int |f_n g| = \liminf_{n\to\infty} \int f_n g \leq \mathcal{M}_q(g). \end{aligned}$$

(For the last estimate we used the remark at the beginning of the proof.) On the other hand, Holder's inequality gives $\mathcal{M}_q(g) \leq ||g||_q$, so the proof is complete for the case $q < \infty$.

Now suppose $q = \infty$. Given $\epsilon > 0$, let $A = \{x : |g(x)| > \mathcal{M}_\infty(g) + \epsilon\}$. If $\mu(A)$ were positive, we could choose $B \subset A$ with $0 < \mu(B) < \infty$ since μ is semifinite. Setting $f = \mu(B)^{-1}\chi_B \overline{\operatorname{sng}\, g}$, we would then have $||f||_1 = 1$, and $\int fg = \mu(B)^{-1}\int_B |g| \geq \mathcal{M}_\infty(g) + \epsilon$. But this is impossible by the remark at the beginning of the proof. Hence $||g||_\infty \leq \mathcal{M}_\infty(g)$, and the reverse inequality is obvious. □

We are now ready to prove the Riesz-Thorin Interpolation Theorem, which allows us to establish boundedness of a linear operator on certain L^p spaces.

Theorem 14.4 (The Riesz-Thorin Interpolation Theorem) *Suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are measure spaces and $p_0, p_1, q_0, q_1 \in [1, \infty]$, $p_0 \leq p_1, q_0 \leq q_1$. If $q_0 = q_1 = \infty$, suppose also that ν is semifinite. For $0 < t < 1$, define p_t and q_t by*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1},\ \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}.$$

If T is a linear map from $L^{p_0}(\mu) + L^{p_1}(\mu)$ into $L^{q_0}(\nu) + L^{q_1}(\nu)$ such that $||Tf||_{q_0} \leq M_0||f||_{p_0}$ for $f \in L^{p_0}(\mu)$ and $||Tf||_{q_1} \leq M_1||f||_{p_1}$ for $f \in L^{p_1}(\mu)$, then $||Tf||_{q_t} \leq M_0^{1-t}M_1^t||f||_{p_t}$ for $f \in L^{p_t}(\mu)$.

Proof. If $p_0 = p_1 = p_t$, then from $Tf \in L^{q_0}(\nu)$, $Tf \in L^{q_1}(\nu)$, we have $Tf \in L^{q_0}(\nu) \cap L^{q_1}(\nu) \subset L^{q_t}$. It follows from Proposition 10.16 that

$$||Tf||_{q_t} \leq ||Tf||_{q_0}^{1-t}||Tf||_{q_1}^t \leq M_0^{1-t}||f||_{p_0}^{1-t}M_1^t||f||_{p_1}^t = M_0^{1-t}M_1^t||f||_{p_t}.$$

Thus we may assume that $p_0 < p_1$, and in particular that $p_t < \infty$ for $0 < t < 1$.

Let $\mathcal{S}_X$ (resp. $\mathcal{S}_Y$) be the space of all simple measurable functions on X (resp. Y) that vanish outside sets of finite measure. Then $\mathcal{S}_X \subset L^p(\mu)$ for all $p \in [1, \infty]$ and $\mathcal{S}_X$ is dense in $L^p(\mu)$ for $1 \leq p < \infty$, similarly for $\mathcal{S}_Y$. The main part of the proof consists of showing

$$||Tf||_{q_t} \leq M_0^{1-t}M_1^t||f||_{p_t},\ \text{for } f \in \mathcal{S}_X. \tag{14.1}$$

Observe that for $f \in \mathcal{S}_X$, we have $f \in L^{p_0}(\mu) \cap L^{p_1}(\mu)$ and so $Tf \in L^{q_0}(\nu) \cap L^{q_1}(\nu) \subset L^{q_t}(\nu)$. Thus if $1 \leq q_0 \leq q_1 < \infty$ or $1 \leq q_0 < q_1 = \infty$, we have

$$\int_Y |Tf|^{q_0} d\nu < \infty,$$

which implies that $\{x : |Tf(x)| > 0\} = \{x : Tf(x) \neq 0\}$ is σ-finite. Also when $q_0 = q_1 = \infty$, we know from the semifiniteness of ν that the conditions in Theorem 14.3 are satisfied. By Theorem 14.3,

$$||Tf||_{q_t} = \sup\left\{\left|\int_Y (Tf)g\ d\nu\right|,\ g \in \mathcal{S}_Y \ and\ ||g||_{q_t'} = 1\right\}, \tag{14.2}$$

where q_t' is the conjugate exponent to q_t. We can assume that $||f||_{p_t} = 1$.

We will establish the following

Claim : $\qquad$ *If* $f \in \mathcal{S}_X$ *and* $||f||_{p_t} = 1$, *then* $\left|\int_Y (Tf)g\ d\nu\right| \leq M_0^{1-t}M_1^t$ *for all* $g \in \mathcal{S}_Y$ *such that* $||g||_{q_t'} = 1$.

Let $f = \sum_{j=1}^m c_j\chi_{E_j}$ and $g = \sum_{k=1}^n d_k\chi_{F_k}$ where the E_j's and the F_k's are disjoint in X and Y and the c_j's and d_k's are nonzero constants. Let $c_j = |c_j|e^{i\theta_j}$ and $d_k = |d_k|e^{i\psi_k}$. Also, let

$$\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1},\ \beta(z) = \frac{1-z}{q_0} + \frac{z}{q_1};$$

then

$$\alpha(t) = \frac{1}{p_t},\ \beta(t) = \frac{1}{q_t},\ 0 < t < 1.$$

Fix $t \in (0,1)$; since $p_t < \infty$, we have $\alpha(t) > 0$. Define

$$f_z = \sum_{j=1}^m |c_j|^{\frac{\alpha(z)}{\alpha(t)}}e^{i\theta_j}\chi_{E_j}.$$

If $\beta(t) < 1$, we define

$$g_z = \sum_{k=1}^n |d_k|^{\frac{(1-\beta(z))}{(1-\beta(t))}}e^{i\psi_k}\chi_{F_k},$$

while if $\beta(t) = 1$ we define $g_z = g$ for all z.

Consider the case that $\beta(t) \neq 1$. Set

$$\begin{aligned}\phi(z) &= \int_Y (Tf_z)g_z d\nu \\ &= \sum_{j,k} e^{i(\theta_j+\psi_k)}|c_j|^{\frac{\alpha(z)}{\alpha(t)}}|d_k|^{\frac{(1-\beta(z))}{(1-\beta(t))}}\int_{F_k}(T\chi_{E_j})d\nu;\end{aligned}$$

then ϕ is a holomorphic function of z that is bounded in the strip $0 \leq Re\ z \leq 1$.

Since $\int(Tf)gd\nu = \phi(t)$, in order to finish the proof of the **Claim**, by the three lines lemma it will suffice to show that $|\phi(z)| \leq M_0$ for $Re\ z = 0$ and $|\phi(z)| \leq M_1$, for $Re\ z = 1$. Now note that since the E_j are disjoint, for any $x \in X$ at most one term of the sum equal to $f(x)$ or $f_z(x)$ may be nonzero. Since

$$\alpha(is) = p_0^{-1} + is(p_1^{-1} - p_0^{-1}),\ 1 - \beta(is) = (1 - q_0^{-1}) - is(q_1^{-1} - q_0^{-1}),\ s \in \mathbb{R},$$

we have, assuming $x \in E_l$, that

$$|f_{is}(x)| = \left||c_l|^{\frac{\alpha(is)}{\alpha(t)}}e^{i\theta_l}\right| = |c_l|^{\frac{p_0^{-1}}{\alpha(t)}} = |c_l|^{\frac{p_t}{p_0}} = |f(x)|^{\frac{p_t}{p_0}}.$$

Assuming that $y \in F_w$, we have

$$|g_{is}(y)| = |d_w|^{\frac{1-q_0^{-1}}{1-\beta(t)}} = |d_w|^{\frac{1-q_0^{-1}}{1-q_t^{-1}}} = |d_w|^{\frac{q_t'}{q_0'}} = |g(y)|^{\frac{q_t'}{q_0'}}$$

Thus

$$\begin{aligned}
|\phi(is)| &= \left|\int (Tf_{is}) g_{is} d\nu\right| \leq ||Tf_{is}||_{q_0} \cdot ||g_{is}||_{q_0'} \leq M_0 ||f_{is}||_{p_0} \cdot ||g_{is}||_{q_0'} \\
&= M_0 \left(\int |f_{is}|^{p_0}\right)^{\frac{1}{p_0}} \left(\int |g_{is}|^{q_0'}\right)^{\frac{1}{q_0'}} \\
&= M_0 \left(\int |f|^{p_t}\right)^{\frac{1}{p_0}} \left(\int |g|^{q_t'}\right)^{\frac{1}{q_0'}} = M_0 ||f||_{p_t}^{p_t/p_0} \cdot ||g||_{q_t'}^{q_t'/q_0} = M_0.
\end{aligned}$$

Similarly, $|\phi(1+is)| \leq M_1$. This completes the proof of the **Claim**. Consequently,

$$||T(f)||_{q_t} \leq M_0^{1-t} M_1^t ||f||_{p_t}, \ \forall f \in \mathcal{S}_X.$$

We have from $p_0 < p_1$ that $p_0 < p_t < p_1$. If $f \in L^{p_t}(\mu)$, take a sequence of simple measurable functions on X such that $|f_n| \leq |f|, \ f_n \to f$ pointwise. Then $||f_n - f||_{p_t} \to 0$ and $\{f_n\} \subset \mathcal{S}_X$. Let $E = \{x : |f(x)| > 1\}, \ g = f\chi_E, \ g_n = f_n \chi_E$, $h = f - g, \ h_n = f_n - g_n$; then $|g|^{p_0} = |f|^{p_0}\chi_E \leq |f|^{p_t}\chi_E$, which implies that $g \in L^{p_0}(\mu)$. Similarly, $h \in L^{p_1}(\mu)$. Observing that $||g_n - g||_{p_0} \to 0$, $||h_n - h||_{p_1} \to 0$, we have $||Tg_n - Tg||_{q_0} \leq M_0||g_n - g_0||_{p_0} \to 0$, $||Th_n - Th||_{q_1} \leq M_1||h_n - h||_{p_1} \to 0$. By passing to a suitable subsequence we may assume that $Tg_n \to Tg$ a.e. and $Th_n \to Th$ a.e.. But then $Tf_n \to Tf$ a.e. From Fatou's lemma, we get

$$\begin{aligned}
\int_Y |Tf|^{q_t} d\nu &\leq \liminf \int_Y |Tf_n|^{q_t} d\nu = \liminf ||Tf_n||_{q_t}^{q_t} \\
&\leq \liminf \left(M_0^{1-t} M_1^t ||f_n||_{p_t}\right)^{q_t} = \left(M_0^{1-t} M_1^t ||f||_{p_t}\right)^{q_t}.
\end{aligned}$$

Therefore,

$$||Tf||_{q_t} \leq M_0^{1-t} M_1^t ||f||_{p_t}, \ \forall f \in L^{p_t}.$$

The case $\beta(t) = 1$ is simpler and left as an exercise. □

We are going to study another interpolation theorem which might be helpful in situations where the Riesz–Thorin interpolation theorem does not apply. In this respect recall, that $f(x) = \frac{1}{x}$ is not integrable over $\mathbb{R}$. To include such functions we begin by slightly weakening the L^p norms. Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to \mathbb{C}$ be a measurable function. We consider the distribution function

$$\mu_f(t) = \mu(\{x \in X : |f(x)| > t\}), \ t \geq 0. \tag{14.3}$$

From Fubini theorem,

$$||f||_p^p = p \int_0^\infty t^{p-1} \mu_f(t) dt, \ 1 \leq p < \infty. \tag{14.4}$$

If $p = \infty$, $||f||_\infty = \inf\{t \geq 0 : \mu_f(t) = 0\}$. We have

$$\begin{aligned}
||f||_p^p &= \int_X |f|^p d\mu \geq \int_{\{x \in X : |f(x)| > t\}} |f|^p d\mu \qquad (14.5) \\
&\geq \int_{\{x \in X : |f(x)| > t\}} t^p d\mu = t^p \mu_f(t),
\end{aligned}$$

and so

$$||f||_p \geq t(\mu_f(t))^{1/p}, \ \forall t > 0.$$

This leads the following definition.

$$||f||_{p,w} = \sup_{t>0} t(\mu_f(t))^{1/p} \le ||f||_p, \tag{14.6}$$

which is called the *weak p-norm of f*. By construction we have

$$||f||_{p,w} \le ||f||_p. \tag{14.7}$$

If $p = \infty$, we set $||\cdot||_{\infty,w} = ||\cdot||_\infty$. Let us denote

$$L^{p,w}(X,\mu) = \{f : X \to \mathbb{C} : ||f||_{p,w} < \infty\}. \tag{14.8}$$

As in the L^p-space, we also regard functions in $L^{p,w}(X,\mu)$ which are equal a.e. as one function. We have seen that $L^p(X,\mu) \subset L^{p,w}(X,\mu)$. It turns out that this inclusion is in general strict.

Example Let $f(x) = \frac{1}{x}$ in $\mathbb{R}$; then $f \notin L^1(\mathbb{R}, dm)$. But

$$\mu_f(t) = m(\{x : |1/x| > t\}) = m(\{x : |x| < \frac{1}{t}\}) = \frac{2}{t}$$

and $||f||_{1,w} = \sup_{t>0} t\mu_f(t) = 2$. Thus $f \in L^{1,w}(\mathbb{R})$. Similarly, for any $p \in [1,\infty)$, the function $g(x) = |x|^{-n/p} \notin L^p(\mathbb{R}^n, dm)$, but $f \in L^{p,w}(\mathbb{R}^n, dm)$. Thus $L^{p,w}(\mathbb{R}^n, dm) \subsetneq L^p(\mathbb{R}^n, dm)$.

Let T be a map from some vector space $\mathcal{D}$ of measurable functions on $(X,\mathcal{A},\mu)$ to the space of all measurable functions on $(Y,\mathcal{B},\nu)$. T is called subadditive if $|T(f+g)| \le |T(f)| + |T(g)|$ and $|T(cf)| = c|T(f)|$ for all $f, g \in \mathcal{D}$ and $c > 0$.

A subadditive operator T is strong type $(p,q)(1 \le p,q \le \infty)$ if $L^p(\mu) \subset \mathcal{D}$, T maps $L^p(\mu)$ into $L^q(\nu)$, and there exists $C > 0$ such that

$$||T(f)||_q \le C_{p,q}||f||_p, \ \forall f \in L^p(\mu).$$

A subadditive operator T is weak type $(p,q)(1 \le p,q < \infty)$ if $L^p(\mu) \subset \mathcal{D}$, T maps $L^p(\mu)$ into $L^{q,w}(\nu)$, and there exists $C_{p,q,w} > 0$ such that

$$||T(f)||_{q,w} \le C_{p,q,w}||f||_p, \ \forall f \in L^p(\mu).$$

Also, we shall say that T is weak type (p,∞) iff T is strong type (p,∞).

By (14.7) strong type (p,q) is stronger than weak type (p,q) and we have $C_{p,q,w} \le C_{p,q}$.

Theorem 14.5 (Marcinkiewicz). *Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be measure spaces and $1 \le p_0 < p_1 \le \infty$. Let T be a subadditive operator defined on $L^p(\mu), p \in [p_0,p_1]$. If T is of weak type (p_0,p_0) and (p_1,p_1) then it is also of strong type (p,p) for every $p_0 < p < p_1$.*

Proof. **Case 1**: $p_1 < \infty$. Fix $f \in L^p(\mu)$, some number $r > 0$ and decompose $f = f_0 + f_1$ according to

$$f_0 = f\chi_{\{x:|f(x)|>r\}} \in L^{p_0} \cap L^p, \ f_1 = f\chi_{\{x:|f(x)|\le r\}} \in L^p \cap L^{p_1}.$$

We have

$$\begin{aligned}
||T(f)||_p^p &= p\int_0^\infty r^{p-1}\nu_{T(f)}(r)dr \\
&= p\int_0^\infty r^{p-1}\nu(\{y\in Y: |T(f)(y)|>r\})dr \\
&= p2^p\int_0^\infty r^{p-1}\nu_{T(f)}(2r)dr.
\end{aligned}$$

It follows from

$$|T(f)| = |T(f_0+f_1)| \le |T(f_0)| + |T(f_1)|$$

that

$$\{y\in Y: |T(f)(y)|>2r\} \subset \{y\in Y: |T(f_0)(y)|>r\} \cup \{y\in Y: |T(f_1)(y)|>r\},$$

and so

$$\nu_{T(f)}(2r) \le \nu_{T(f_0)}(r) + \nu_{T(f_1)}(r).$$

Since T is of weak type (p_0,p_0), we have

$$||T(f_0)||_{p_0,w} = \sup_{t>0} t(\nu_{T(f_0)}(t))^{1/p_0} \le C_0||f_0||_{p_0}.$$

Therefore,

$$\nu_{T(f_0)}(r) \le \left(\frac{C_0}{r}\right)^{p_0}\cdot ||f_0||_{p_0}^{p_0},\ \forall r>0.$$

Similarly,

$$\nu_{T(f_1)}(r) \le \left(\frac{C_1}{r}\right)^{p_1}\cdot ||f_1||_{p_1}^{p_1},\ \forall r>0.$$

Thus we have

$$\begin{aligned}
||T(f)||_p^p &\le p2^p\left(\int_0^\infty r^{p-1}\nu_{T(f_0)}(r)dr + \int_0^\infty r^{p-1}\nu_{T(f_1)}(r)dr\right) \\
&\le p2^p\left(\int_0^\infty r^{p-1}\left(\frac{C_0}{r}\right)^{p_0}\cdot||f_0||_{p_0}^{p_0} + \int_0^\infty r^{p-1}\left(\frac{C_1}{r}\right)^{p_1}\cdot||f_1||_{p_1}^{p_1}\right)dr
\end{aligned}$$

Observe that

$$\begin{aligned}
||f_0||_{p_0}^{p_0} &= \int_X |f_0|^{p_0}d\mu = \int_X |f|^{p_0}\chi_{\{x\in X:|f(x)|>r\}}d\mu \\
&= \int_{\{x\in X:|f(x)|>r\}} |f|^{p_0}d\mu
\end{aligned}$$

and

$$||f_1||_{p_1}^{p_1} = \int_{\{x\in X:|f(x)|\le r\}} |f|^{p_1}d\mu.$$

Hence

$$||T(f)||_p^p \le p2^p(C_0^{p_0}I_0 + C_1^{p_1}I_1),$$

where

$$\begin{aligned} I_0 &= \int_0^\infty \left[\int_{\{x\in X:|f(x)|>r\}} r^{p-p_0-1}|f|^{p_0}d\mu\right]dr \qquad (14.9)\\ &= \int_0^\infty \left[\int_X r^{p-p_0-1}|f|^{p_0}\chi_{\{(x,r)\in X\times\mathbb{R}^+:|f(x)|>r\}}d\mu\right]dr\\ &= \int_X \left[\int_0^\infty r^{p-p_0-1}|f|^{p_0}\chi_{\{(x,r)\in X\times\mathbb{R}^+:|f(x)|>r\}}dr\right]d\mu\\ &= \int_X \left[\int_0^{|f|} r^{p-p_0-1}|f|^{p_0}dr\right]d\mu = \frac{1}{p-p_0}||f||_p^p, \end{aligned}$$

and

$$\begin{aligned} I_1 &= \int_0^\infty \left[\int_{\{x\in X:|f(x)|\le r\}} r^{p-p_1-1}|f|^{p_1}d\mu\right]dr \qquad (14.10)\\ &= \int_X \left[\int_{|f|}^\infty r^{p-p_1-1}|f|^{p_1}dr\right]d\mu = \frac{1}{p_1-p}||f||_p^p. \end{aligned}$$

Hence,

$$||T(f)||_p \le C||f||_p.$$

Case 2: $p_1 = \infty$.

In this case, we have

$$||T(g)||_\infty \le C_1||g||_\infty,\ \forall g \in L^\infty(\mu)$$

and

$$||T(h)||_{p_0,w} \le C_0||h||_{p_0},\ \forall h \in L^{p_0}(\mu).$$

The case $C_1 = 0$ can be settled easily. We assume that $C_1 > 0$. Let $f \in L^p(\mu)$ and set

$$f_0 = f\chi_{\{x\in X:|f(x)|>\frac{r}{C_1}\}} \in L^{p_0}\cap L^p,$$

$$f_1 = f\chi_{\{x\in X:|f(x)|\le\frac{r}{C_1}\}} \in L^p\cap L^\infty.$$

Since $||f_1||_\infty \le r/C_1$, we have

$$||T(f_1)||_\infty \le r,$$

which implies that

$$\nu_{T(f_1)}(r) = \nu(\{y\in Y : |T(f_1)| > r\}) = 0.$$

Hence

$$\begin{aligned} \nu_{T(f)}(2r) &\le \nu_{T(f_0)}(r) + \nu_{T(f_1)}(r) = \nu_{T(f_0)}(r)\\ &\le \frac{C_0^{p_0}}{r^{p_0}}||f_0||_{p_0}^{p_0} = \frac{C_0^{p_0}}{r^{p_0}}\int_{\{x\in X:|f(x)|>\frac{r}{C_1}\}}|f|^{p_0}d\mu. \end{aligned}$$

It then follows from the same calculations as in the **Case 1** that

$$\begin{aligned}
||T(f)||_p^p &= p2^p\int_0^\infty r^{p-1}\nu_{T(f)}(2r)dr\\
&\le p2^pC_0^{p_0}\int_0^\infty\left[\int_{\{x\in X:|f(x)|>\frac{r}{C_1}\}} r^{p-p_0-1}|f|^{p_0}d\mu\right]dr\\
&= p2^pC_0^{p_0}\int_X|f(x)|^{p_0}\left[\int_0^{C_1|f(x)|} r^{p-p_0-1}dr\right]d\mu\\
&= \frac{p2^pC_0^{p_0}C_1^{p-p_0}}{p-p_0}\int_X|f|^pd\mu.
\end{aligned}$$

Therefore,

$$||T(f)||_p\le 2\left(\frac{p}{p-p_0}\right)^{1/p}C_0^{p_0/p}C_1^{1-p_0/p}||f||_p.$$

□

Exercises 13

1. Let $(X,\mathcal{A},\mu)$ be a σ-finite measure space, f a measurable function on X, and $0<r<p<\infty$. Prove that $f\in L^{p,w}(\mu)$ iff there exists a positive finite constant c such that $\int_X|f|^rd\mu\le c\mu(A)^{1-r/p}$ for every measurable A with $\mu(A)<\infty$.

2. If f is a nonnegative measurable function defined on the positive real numbers, let

$$F(x)=\frac{1}{x}\int_0^x f(t)dt,\ G(x)=\frac{1}{x}\int_x^\infty f(t)dt.$$

Let $1\le p<\infty, r>0$. Show that

$$\int_0^\infty (F(x))^px^{p-1-r}dx\le\left(\frac{p}{r}\right)^p\int_0^\infty (f(x))^px^{p-1-r}dx$$

and

$$\int_0^\infty (G(x))^px^{p-1+r}dx\le\left(\frac{p}{r}\right)^p\int_0^\infty (f(x))^px^{p-1+r}dx.$$

15. Fractional integral and Hardy-Littlewood-Sobolev inequality

The purpose of this section is to prove the Hardy-Little-Wood Sobolev inequality. Firstly we need some preliminaries.

If f and g are measurable in $\mathbb{R}^n$, their convolution $(f*g)(x)$ is defined by

$$(f*g)(x)=\int_{\mathbb{R}^n}f(x-t)g(t)dt,$$

provided the integral exists.

We first claim that $f*g=g*f$, that is,

$$\int_{\mathbb{R}^n}f(x-t)g(t)dt=\int_{\mathbb{R}^n}f(t)g(x-t)dt. \tag{15.1}$$

This follows from changes of variable. In our case, it amounts to the statement that if $x \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} r(t)dt = \int_{\mathbb{R}^n} r(x-t)dt \tag{15.2}$$

when $r(t) = f(x-t)g(t)$. For fixed x, the map $T : t \to x - t$ is an isometry from $\mathbb{R}^n$ to $\mathbb{R}^n$. Therefore, (15.2) holds.

A map $T : \mathbb{R}^n \to \mathbb{R}^n$ is called a Lipschitz transformation if there is a constant c such that

$$|Tx - Ty| \leq c|x - y|, \ \forall x, y \in \mathbb{R}^n.$$

The smallest such constant c, namely, the number

$$c = \sup_{x,y \in \mathbb{R}^n, x \neq y} \frac{|Tx - Ty|}{|x - y|}$$

is called the *Lipschitz constant* of T.

An n-dimensional interval I is a subset of $\mathbb{R}^n$ of the form $I = \{x = (x_1, ..., x_n) : a_k \leq x_k \leq b_k, k = 1, ..., n\}$, where $a_k < b_k, k = 1, ..., n$. An interval is thus closed, and we say it has edges parallel to the coordinate axes. If the edge lengths $b_k - a_k$ are all equal, I will be called an n-dimensional cube with edges parallel to the coordinate axes. A set equal to an interval minus some part of its boundary will be called a partly open interval.

It is known that every open set in $\mathbb{R}$ can be written as a countable union of disjoint open intervals. In the high dimensional case, every open set in $\mathbb{R}^n$, $n \geq 1$, can be written as a countable union of nonoverlapping (closed) cubes. It can also be written as a countable union of disjoint partly open cubes.

A continuous map may not preserve measurability. However, there holds the following

Theorem 15.1 *If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a Lipschitz map, then T maps measurable sets into measurable sets.*

Proof. We will first show that a continuous transformation sends sets of type F_σ into sets of type F_σ. A continuous T maps compact sets into compact sets; therefore, since any closed set can be written as a countable union of compact sets, T maps closed sets into sets of type F_σ. Here, we have used the relation $T(\bigcup_{i=1}^{\infty} E_i) = \bigcup_{i=1}^{\infty} TE_i$, which holds for any T and $\{E_k\}$. It follows that a continuous T preserves the class of F_σ sets.

We will next show that a Lipschitz map T sends sets of measure zero into sets of measure zero. Since $|Tx - Tx| \leq c|x - x|$, the image of a set with diameter d has diameter at most cd. It follows by covering with cubes that there is a constant c' such that $|TI| \leq c'|I|$ for any interval I; note that TI is measurable since I is closed. If $|Z| = 0$ and $\epsilon > 0$, choose intervals $\{I_k\}$ covering Z such that $\sum_k |I_k| < \epsilon$. Since $TZ \subset \cup_k TI_k$, we have $\mu^*(TZ) \leq \sum_k |TI_k| \leq c' \sum_k |I_k| \leq c'\epsilon$. Hence, TZ is a set of measure zero.

If E is a measurable set, we can write $E = H \cup Z$, where H is of type F_σ and $|Z| = 0$. Since $TE = TH \cup TZ$, TE is measurable as the union of measurable sets.

This completes the proof. $\square$

Lemma 15.2 *If $f(x)$ is measurable in $\mathbb{R}^n$, then the function $F(x,t) = f(x-t)$ is measurable in $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$.*

Proof. Let $F_1(x,t) = f(x)$. Since f is measurable, it follows that $F_1(x,t)$ is measurable in $\mathbb{R}^{2n}$. In fact, the set $\{(x,t) : F_1(x,t) > a\}$, equals $\{(x,t) : f(x) > a, t \in \mathbb{R}^n\} = \{x : f(x) > a\} \times \mathbb{R}^n$ and so is measurable in $\mathbb{R}^n \times \mathbb{R}^n$ with the product measure.

For $(x,t) \in \mathbb{R}^{2n}$, consider the transformation $T(x,t) = (x-t, x+t)$. This is a nonsingular linear transformation of $\mathbb{R}^{2n}$, which implies that T^{-1} is Lipschitz. Therefore, by Theorem 15.1, the function F_2 defined by $F_2(x,t) = F_1(x-t, x+t) = F_1 \circ T(x,t)$ is measurable in $\mathbb{R}^{2n}$. Since $F_2(x,t) = f(x-t)$, the lemma follows. $\square$

Theorem 15.3 *(1) If $f, g \in L(\mathbb{R}^n) \equiv L^1(\mathbb{R}^n, m)$, then $(f * g)(x)$ exists for almost every $x \in \mathbb{R}^n$ and is measurable. Moreover, $f * g \in L(\mathbb{R}^n)$ and*

$$\int_{\mathbb{R}^n} |f * g| dx \le \left(\int_{\mathbb{R}^n} |f| dx\right)\left(\int_{\mathbb{R}^n} |g| dx\right),$$

$$\int_{\mathbb{R}^n} f * g dx = \left(\int_{\mathbb{R}^n} f dx\right)\left(\int_{\mathbb{R}^n} g dx\right).$$

(2) If $1 < p \le \infty, f \in L^1$, and $g \in L^p$, then $||f * g||_p \le ||f||_1 ||g||_p$.

Proof. Suppose first that both f and g are nonnegative on $\mathbb{R}^n$. By Lemma 15.2, $f(x-t)g(t)$ is measurable on $\mathbb{R}^n \times \mathbb{R}^n$. Hence, the integral

$$I = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-t)g(t) dt dx$$

is well defined. By Tonelli's theorem and (15.2),

$$\begin{aligned} I &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-t)g(t)dt\right] dx \\ &= \int_{\mathbb{R}^n} g(t) \left[\int_{\mathbb{R}^n} f(x-t)dx\right] dt = \left[\int_{\mathbb{R}^n} g(t)dt\right]\left[\int_{\mathbb{R}^n} f(x)dx\right]. \end{aligned}$$

The first of these equations can be written $I = \int_{\mathbb{R}^n} (f*g)(x)dx$, where measurability of $f * g$ is guaranteed by Tonelli's theorem. This proves (1) for $f \ge 0$ and $g \ge 0$. For general $f, g \in L^1(\mathbb{R}^n)$, it follows that

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-t)||g(t)| dt dx = \int_{\mathbb{R}^n} (|f| * |g|) dx = \left[\int_{\mathbb{R}^n} |f| dt\right]\left[\int_{\mathbb{R}^n} |g| dx\right] < \infty.$$

Hence $f(x-t)g(t)$ is integrable in $\mathbb{R}^{2n}$. By Fubini's theorem, $\int_{\mathbb{R}^n} f(x-t)g(t)dt$ exists for a.e. x and is measurable and integrable; also,

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-t)g(t) dt dx = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-t)g(t)dt\right] dx = \left[\int_{\mathbb{R}^n} f dt\right]\left[\int_{\mathbb{R}^n} g dx\right]$$

Thus
$$\int_{\mathbb{R}^n} f * g dx = \left[\int_{\mathbb{R}^n} f dt\right]\left[\int_{\mathbb{R}^n} g dx\right].$$
Finally, since $|f * g| \leq |f| * |g|$ wherever $f * g$ exists, we obtain by integration that
$$\int_{\mathbb{R}^n} |f * g| \leq \int_{\mathbb{R}^n} (|f| * |g|) = \left[\int_{\mathbb{R}^n} |f| dt\right]\left[\int_{\mathbb{R}^n} |g| dx\right],$$
This is completes the proof of (1).

(2) The case $p = \infty$ is easy and left as an exercise, so let us suppose $p < \infty$. Let q be the conjugate exponent to p. By Hölder's inequality
$$\begin{aligned} & \left|\int_{\mathbb{R}^n} f(y)g(x-y)dy\right| \\ \leq & \int_{\mathbb{R}^n} |f(y)|^{\frac{1}{q}}|f(y)|^{1-\frac{1}{q}}|g(x-y)|dy \\ \leq & \left(\int_{\mathbb{R}^n} |f(y)|dy\right)^{1/q}\left(\int_{\mathbb{R}^n} |f(y)|^{p(1-\frac{1}{q})}|g(x-y)|^p dy\right)^{1/p}. \end{aligned}$$
Then, using the Fubini theorem,
$$\begin{aligned} \int_{\mathbb{R}^n} |f * g(x)|^p dx & \leq \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(y)|dy\right)^{p/q}\int_{\mathbb{R}^n} |f(y)||g(x-y)|^p dy dx \\ & = ||f||_1^{p/q}||g||_p^p \int_{\mathbb{R}^n} |f(y)|dy \\ & = ||f||_1^{1+p/q}||g||_p^p. \end{aligned}$$
Taking p^{th} roots gives our desired result. □

In the proof of the above Theorem, we showed the following useful fact.

Corollary 15.4 If f and g are nonnegative and measurable on $\mathbb{R}^n$, then $f * g$ is measurable on $\mathbb{R}^n$ and
$$\int_{\mathbb{R}^n} f * g dx = \left(\int_{\mathbb{R}^n} f dx\right)\left(\int_{\mathbb{R}^n} g dx\right).$$

Let f be a real-valued measurable function on $\mathbb{R}^n$, $n \geq 1$, and let $0 < \alpha < n$. The fractional integral or Riesz potential of f of order α is defined by
$$I_\alpha f(x) = \int_{\mathbb{R}^n} f(y)\frac{1}{|x-y|^{n-\alpha}}dy,\ x \in \mathbb{R}^n. \tag{15.3}$$
provided the integral exists. By allowing f to vary, the mapping defined by $I_\alpha : f \to I_\alpha f$, that is, the convolution operator with kernel $|x|^{\alpha-n}$, is called the *fractional integral operator of order* α. Note that if f is nonnegative and measurable on $\mathbb{R}^n$, then since $|x|^{\alpha-n}$ is also nonnegative and measurable, Corollary 15.4 guarantees existence and measurability (but not finiteness) of $I_\alpha f$ on $\mathbb{R}^n$.

We will now determine values of p and q for which I_α is a bounded operator from $L^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$ and study some closely related "endpoint" estimates. We always assume $0 < \alpha < n$ and use the notation $||f||_p$ for the $L^p(\mathbb{R}^n)$ norm of $f, 1 \leq p \leq \infty$.

It will turn out that the values of p and q for which the norm inequality

$$||I_\alpha f||_q \leq c||f||_p \ for\ all\ f \in L^p(\mathbb{R}^n) \tag{15.4}$$

is true, for some c independent of f, are limited to

$$1 < p < \frac{n}{\alpha} \text{ and } \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}.$$

Let us list three comments that explain why the restrictions on p and q just mentioned are necessary for (15.4).

i) If $p \geq \frac{n}{\alpha}$, there exists $f \in L^p(\mathbb{R}^n)$ such that $I_\alpha f = \infty$ everywhere in $\mathbb{R}^n$. In particular, (15.4) cannot hold if $p \geq \frac{n}{\alpha}$.

ii) If $p = 1$ and $\frac{1}{q} = 1 - \frac{\alpha}{n}$, there exists $f \in L^1(\mathbb{R}^n)$such that $||I_\alpha f||_q = \infty$. Thus, (15.4) fails when $p = 1$ and $q = n/(n-\alpha)$.

iii) If $1 \leq p < \frac{n}{\alpha}$, the only value of q for which (15.4) can possibly hold for all $f \in L^p(\mathbb{R}^n)$ satisfies $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$, that is, $q = pn/(n-\alpha)$.

To verify i), let α, p satisfy $0 < \alpha < n$ and $n/\alpha \leq p \leq \infty$. Let ϕ denote the characteristic function of $\{y : |y| > 2\}$, and let

$$f(y) = \frac{\phi(y)}{|y|^\alpha \log |y|}.$$

Then f is bounded and nonnegative on $\mathbb{R}^n$, and also $f \in L^p(\mathbb{R}^n)$ for $\frac{n}{\alpha} \leq p < \infty$ since

$$\int_{\mathbb{R}^n} f^p = \int_{\{|y|>2\}} \frac{1}{|y|^{\alpha p}(\log |y|)^p} = c_{n-1} \int_2^\infty t^{n-1} \frac{1}{t^{\alpha p}(\log t)^p} dt < \infty,$$

where c_{n-1} denotes the area of the unit sphere $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$ and we have used the fact that $p > 1$ in case $p = n/\alpha$. However, $I_\alpha f(x) = \infty$ for every x since

$$\begin{aligned}
I_\alpha f(x) &= \int_{\mathbb{R}^n} \frac{\phi(y)}{|y|^\alpha \log |y|} \frac{1}{|x-y|^{n-\alpha}} dy \\
&\geq \int_{\{|y|>2+|x|\}} \frac{1}{|y|^\alpha \log |y|} \frac{1}{|x-y|^{n-\alpha}} dy \\
&\geq \int_{\{|y|>2+|x|\}} \frac{1}{|y|^\alpha \log |y|} \frac{1}{(2|y|)^{n-\alpha}} dy \\
&= 2^{\alpha-n} \int_{\{|y|>2+|x|\}} \frac{1}{|y|^n \log |y|} dy = \infty.
\end{aligned}$$

In case $p = \infty$, a much simpler function f can be chosen above; for example, the constant function $f \equiv 1$ has fractional integral equal to

$$\int_{\mathbb{R}^n} \frac{1}{|x-y|^{n-\alpha}} dy = \int_{\mathbb{R}^n} \frac{1}{|y|^{n-\alpha}} dy = \infty.$$

To verify ii), let $0 < \alpha < n$ and $p = 1$. If B denotes the unit ball in $\mathbb{R}^n$, then $\chi_B \in L^1(\mathbb{R}^n)$ and

$$\begin{aligned} I_\alpha \chi_B(x) &= \int_{\{|y|<1\}} \frac{1}{|x-y|^{n-\alpha}} dy \\ &\geq \int_{\{|y|<1\}} \frac{1}{(1+|x|)^{n-\alpha}} dy \\ &= \frac{c_{n-1}}{n(1+|x|)^{n-\alpha}}, \; x \in \mathbb{R}^n. \end{aligned}$$

Hence, if $q = n/(n-\alpha)$, then

$$\int_{\mathbb{R}^n} (I_\alpha \chi_B(x))^q dx \geq \left(\frac{c_{n-1}}{n}\right)^q \int_{\mathbb{R}^n} \frac{1}{(1+|x|)^n} dx = \infty.$$

Finally we verify (iii). For $\lambda \in (0, \infty)$, we denote $\delta_\lambda f$ the dilation of f defined by $\delta_\lambda f(x) = f(\lambda x)$. A computation shows that $\delta_\lambda(I_\alpha f) = \lambda^\alpha I_\alpha(\delta_\lambda f)$. The rest is left as an exercise.

The next result gives basic norm estimates for the Riesz operators I_α.

Theorem 15.5 (Hardy-Littlewood-Sobolev) *Let*

$$0 < \alpha < n, \; 1 \leq p < \frac{n}{\alpha}, \; \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}.$$

Then for every $f \in L^p(\mathbb{R}^n)$, $I_\alpha f$ is measurable in $\mathbb{R}^n$. Moreover,

a) if $1 < p < \frac{n}{\alpha}$, then

$$||I_\alpha f||_q \leq c||f||_p$$

for a constant c that depends only on α, n, and p;

b) if $p = 1$, then

$$\sup_{\lambda > 0} \lambda \left|\{x \in \mathbb{R}^n : |I_\alpha f(x)| > \lambda\}\right|^{1/q} \leq c||f||_1 \; \left(q = \frac{n}{n-\alpha}\right).$$

Proof. Let f be nonnegative and measurable on $\mathbb{R}^n$. Then, as noted before, $I_\alpha f$ is measurable on $\mathbb{R}^n$. For $\delta > 0$ to be chosen and $x \in \mathbb{R}^n$, we write

$$I_\alpha f(x) = \int_{|x-y|<\delta} \frac{f(y)}{|x-y|^{n-\alpha}} dy + \int_{|x-y|\geq\delta} \frac{f(y)}{|x-y|^{n-\alpha}} dy \equiv J_1(x) + J_2(x).$$

If $1 < p < \frac{n}{\alpha}$, $\frac{1}{p} + \frac{1}{p'} = 1$, then

$$J_2(x) \leq ||f||_p \left(\int_{|x-y|\geq\delta} \frac{1}{|x-y|^{(n-\alpha)p'}} dy\right)^{1/p'} = ||f||_p C_{n,\alpha,p} \delta^{\alpha - \frac{n}{p}},$$

since $(n-\alpha)p' > n$ due to the condition $p < n/\alpha$. In case $p = 1$, so that $p' = \infty$, then

$$J_2(x) = \int_{|x-y|\geq\delta} \frac{f(y)}{|x-y|^{n-\alpha}} dy \leq \delta^{\alpha-n} \int_{\mathbb{R}^n} f(y) dy = \delta^{\alpha-n} ||f||_1.$$

Next, we will show that $J_1(x) \leq c_{n,\alpha} \delta^\alpha f^*(x)$, where f^* denotes the Hardy–Littlewood maximal function of f. That is,

$$f^*(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y) dy.$$

We have

$$\begin{aligned} J_1(x) &= \sum_{i=1}^{\infty} \int_{\delta 2^{-i} \le |x-y| < \delta 2^{-i+1}} \frac{f(y)}{|x-y|^{n-\alpha}} dy \\ &\le \sum_{i=1}^{\infty} \frac{1}{(\delta 2^{-i})^{n-\alpha}} \int_{|x-y|<\delta 2^{-i+1}} f(y) dy \\ &\le d_n \sum_{i=1}^{\infty} \frac{(\delta 2^{-i+1})^n}{(\delta 2^{-i})^{n-\alpha}} f^*(x) = c_{n,\alpha} \delta^{\alpha} f^*(x). \end{aligned}$$

By combining the estimates for J_1 and J_2, we obtain

$$I_\alpha f(x) \le c \left[\delta^\alpha f^*(x) + \delta^{\alpha - \frac{n}{p}} ||f||_p \right], \tag{15.5}$$

where the constant c depends only on n, α, and p. If $f^*(x) = 0$ for some $x \in \mathbb{R}^n$, then $f = 0$ a.e. in $\mathbb{R}^n$ and conclusions (a) and (b) are trivially true. Thus we can assume that $f^*(x) \neq 0$, for any $x \in \mathbb{R}^n$. When $f^*(x) \neq \infty$, choosing δ (depending on f and x) such that the two terms on the right side are the same, namely, $\delta = (||f||_p / f^*(x))^{p/n}$, gives

$$I_\alpha f(x) \le c ||f||_p^{\frac{\alpha p}{n}} f^*(x)^{1 - \frac{\alpha p}{n}}. \tag{15.6}$$

On the other hand, if $f^*(x) = \infty$, then (15.6) is automatically true and no choice of δ is needed.

It follows from (15.6) that

$$\begin{aligned} ||I_\alpha f||_q &\le c ||f||_p^{\frac{\alpha p}{n}} \left(\int_{\mathbb{R}^n} f^*(x)^{q(1-\frac{\alpha p}{n})} \right)^{1/q} \qquad (15.7) \\ &= c ||f||_p^{\frac{\alpha p}{n}} \left(\int_{\mathbb{R}^n} f^*(x)^p \right)^{1/q} \\ &\le c' ||f||_p^{\frac{\alpha p}{n} + \frac{p}{q}} = c' ||f||_p, \end{aligned}$$

where in order to obtain the last inequality, we have assumed $p > 1$ and applied Theorem 13.8. Note that the constant c' still depends only on n, α, and p. This completes the proof of part (a) when $f \ge 0$.

If $p = 1$, then (15.6) implies that for any $\lambda > 0$,

$$\{x \in \mathbb{R}^n : I_\alpha f(x) > \lambda\} \subset \left\{ x \in \mathbb{R}^n : f^*(x) > \left(\frac{\lambda}{c||f||_1^{\alpha/n}} \right)^{\frac{n}{n-\alpha}} \right\}, \tag{15.8}$$

assuming as we may that $||f||_1 \neq 0$. Applying Theorem 13.5 to the set on the right side yields

$$|\{x \in \mathbb{R}^n : I_\alpha f(x) > \lambda\}| \le c_1 \left(\frac{c||f||_1^{\alpha/n}}{\lambda} \right)^{\frac{n}{n-\alpha}} ||f||_1 = C \left(\frac{||f||_1}{\lambda} \right)^{\frac{n}{n-\alpha}},$$

with C depending only on n and α. This agrees with the inequality in part (b). The theorem is proved for nonnegative f.

Next, consider a general $f \in L^p(\mathbb{R}^n)$ for $1 \le p < n/\alpha$ and $0 < \alpha < n$. The results just derived for nonnegative measurable functions can be applied to $|f|$ to conclude that (a) and (b) hold with $I_\alpha(|f|)$ in place of $I_\alpha f$. Thus we have $||I_\alpha(|f|)||_q \le C|||f|||_p = C||f||_p < \infty$ and so $I_\alpha(|f|)$ is finite a.e., which implies that

$I_\alpha f$ is finite a.e.. Observing that $|I_\alpha f(x)| \le I_\alpha(|f|)(x)$ at any point x where $I_\alpha f(x)$ exists, we conclude that $||I_\alpha f||_q \le ||I_\alpha(|f|)||_q \le C||f||_p$. The proof of Theorem is now complete.

□

Exercises 14

1. a) Verify the dilation property $\delta_\lambda(I_\alpha f) = \lambda^\alpha I_\alpha(\delta_\lambda f)$, where $\delta_\lambda f(x) = f(\lambda x),\ \lambda > 0$.

b) Show that

$$\frac{||I_\alpha f||_q}{||f||_p} = \lambda^{\alpha+\frac{n}{q}-\frac{n}{p}} \frac{||I_\alpha(\delta_\lambda f)||_q}{||\delta_\lambda f||_p},\ \forall \lambda > 0,$$

c) Show that (iii) holds.

16. The Fourier Transform

The Fourier transform $\hat{f}(x)$ of a function f on $\mathbb{R}^n$, $n \ge 1$, defined formally as

$$\hat{f}(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(y) e^{-ix\cdot y} dy. \tag{16.1}$$

Both f and $\hat{f}$ may be complex-valued. One of our main goals is to prove that the mapping $f \to \hat{f}$ is essentially an isometry on $L^2(\mathbb{R}^n)$. An important requirement for achieving this is to find an interpretation of $\hat{f}$ in case $f \in L^2(\mathbb{R}^n)$. The integral in (16.1) may not converge absolutely for every $f \in L^2(\mathbb{R}^n)$. However, as is easy to see, (16.1) does converge absolutely if $f \in L^1(\mathbb{R}^n)$. Properties of $\hat{f}$ when $f \in L^1(\mathbb{R}^n)$ are simpler to derive precisely because of this absolute convergence, and we begin with that case. Furthermore, properties of $\hat{f}$ when $f \in L^1(\mathbb{R}^n)$ will be useful later for studying $\hat{f}$ for other classes of functions f.

16.1. The Fourier Transform on L^1. In this subsection, we list some properties of the Fourier transform of functions in $L^1(\mathbb{R}^n)$.

(1) Let $f \in L^1(\mathbb{R}^n)$ and define $\hat{f}(x)$ by (16.1). Note that

$$|\hat{f}(x)| = \left| \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(y) e^{-ix\cdot y} dy \right| \le \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |f(y)| dy,\ x \in \mathbb{R}^n.$$

Thus, the integral in (16.1) converges absolutely and so exists for every $f \in L^1(\mathbb{R}^n)$ and every $x \in \mathbb{R}^n$, and the mapping $f \to \hat{f}$ sends the space $L^1(\mathbb{R}^n)$ into $L^\infty(\mathbb{R}^n)$ with

$$\sup_{x\in\mathbb{R}^n} |\hat{f}(x)| \le (2\pi)^{-n} ||f||_1. \tag{16.2}$$

(2) The mapping $f \to \hat{f}$ is linear on $L^1(\mathbb{R}^n)$.

(3) $\hat{f}(x)$ is a uniformly continuous function of x on $\mathbb{R}^n$ if $f \in L^1(\mathbb{R}^n)$. In fact, for any $x, h \in \mathbb{R}^n$,

$$\begin{aligned}(2\pi)^n|\hat{f}(x+h) - \hat{f}(x)| &= \left|\int_{\mathbb{R}^n} f(y)e^{-ix\cdot y}\{e^{-iy\cdot h} - 1\}dy\right| \\ &\le \int_{\mathbb{R}^n} |f(y)| \min\{|h||y|, 2\}dy \\ &\le \int_{|y|<N} |f(y)||h||y|dy + \int_{|y|\ge N} |f(y)|2dy \\ &= I + II,\end{aligned}$$

where $N > 0$ will be chosen soon. Let $\epsilon > 0$ and note that II depends only on N and f, not on x or h, and $II \to 0$ as $N \to \infty$ since $f \in L^1(\mathbb{R}^n)$. Fix N so large that $II < \epsilon/2$. For I, we have

$$I \le \int_{|y|<N} |f(y)||h||y|dy \le ||f||_1 N|h| < \epsilon/2$$

if $|h| < \epsilon/(2N||f||_1)$. Here, we have assumed that $||f||_1 \ne 0$, but otherwise $\hat{f} \equiv 0$ and the result is trivial. Hence, $I + II < \epsilon$ uniformly in x and h provided $|h|$ is small, which proves the result.

(4) The following is known as the *Riemann–Lebesgue Theorem*.

Theorem 16.1 (Riemann–Lebesgue) *If $f \in L^1(\mathbb{R}^n)$, then*

$$\lim_{|x|\to\infty} \hat{f}(x) = 0. \tag{16.3}$$

Proof. For any $x \in \mathbb{R}^n, |x| \ne 0$, we write

$$\begin{aligned}(2\pi)^n\hat{f}(x) &= \int_{\mathbb{R}^n} f(y)e^{-ix\cdot y}dy \\ &= \int_{\mathbb{R}^n} f\left(y + \frac{\pi x}{|x|^2}\right) e^{-ix\cdot y}e^{-i\pi\frac{|x|^2}{|x|^2}}dy = -\int_{\mathbb{R}^n} f\left(y + \frac{\pi x}{|x|^2}\right) e^{-ix\cdot y}dy.\end{aligned}$$

Adding the first and last formulas gives

$$2(2\pi)^n\hat{f}(x) = \int_{\mathbb{R}^n} \left[f(y) - f\left(y + \frac{\pi x}{|x|^2}\right)\right] e^{-ix\cdot y}dy$$

and so

$$2(2\pi)^n|\hat{f}(x)| \le \int_{\mathbb{R}^n} \left|\left[f(y) - f\left(y + \frac{\pi x}{|x|^2}\right)\right]\right| dy$$

The last integral tends to 0 as $|x| \to \infty$ by continuity in L^1, Theorem 11.12, and the proof is complete. □

(5) (Translation) Let $f \in L^1(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$. Define the *translation* $\tau_h f$ of f by h as $(\tau_h f)(x) = f(x + h)$. Then

$$\widehat{(\tau_h f)}(x) = e^{ix\cdot h}\hat{f}(x), \; x \in \mathbb{R}^n. \tag{16.4}$$

Also,

$$(\tau_h \hat{f})(x) = \hat{f}(x+h) = \widehat{E_h f}(x),$$

where $E_h f(z) = f(z)e^{-iz\cdot h}$

(6) (Dilation) Let $f \in L^1(\mathbb{R}^n)$ and $\lambda \in \mathbb{R} - \{0\}$. Define the dilation $\delta_\lambda f$ of f by λ to be the function

$$\delta_\lambda f(x) = f(\lambda x).$$

Then

$$\widehat{\delta_\lambda f}(x) = \frac{1}{|\lambda|^n}\hat{f}\left(\frac{x}{\lambda}\right),\ x \in \mathbb{R}^n. \tag{16.5}$$

(7) (Rotation) Let $\mathcal{O}$ be an orthogonal linear transformation of $\mathbb{R}^n$ and set $(\mathcal{O}f)(x) = f(\mathcal{O}x),\ x \in \mathbb{R}^n$. If $f \in L^1(\mathbb{R}^n)$, then

$$\widehat{\mathcal{O}f}(x) = \mathcal{O}\hat{f}(x)(= \hat{f}(\mathcal{O}x)),\ x \in \mathbb{R}^n.$$

In fact,

$$\begin{aligned}
\widehat{\mathcal{O}f}(x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \mathcal{O}f(y)e^{-ix\cdot y}dy \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(\mathcal{O}y)e^{-ix\cdot y}dy \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(y)e^{-ix\cdot\mathcal{O}^{-1}y}\frac{1}{|\det\,\mathcal{O}|}dy \\
&= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(y)e^{-i\mathcal{O}x\cdot y}dy = \hat{f}(\mathcal{O}x).
\end{aligned}$$

By definition, a radial function of x is one that depends only on $|x|$. Thus, f is a radial function on $\mathbb{R}^n$ if there is a function $g(t), t \geq 0$, such that $f(x) = g(|x|)$ for all $x \in \mathbb{R}^n$. Also, f is a radial function on $\mathbb{R}^n$ if it is invariant under rotation about the origin. Formula (16.4) says that rotations about the origin commute with the Fourier transform. As a consequence, we obtain the next property.

(8) The Fourier transform of an integrable radial function is a radial function.

(9) Let $f(x) = e^{-|x|^2},\ x \in \mathbb{R}^n$. Then

$$\hat{f}(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-|y|^2}e^{-ix\cdot y}dy = \frac{1}{(2\sqrt{\pi})^n}e^{-|x|^2/4}. \tag{16.6}$$

Proof. We consider first the case that $n = 1$. Let $g(x) = \int_{\mathbb{R}} e^{-|y|^2}e^{-ixy}dy$. We may differentiate with respect to x under the integral sign because $(e^{-i(x+h)y} - e^{-ixy})/h$ is bounded in absolute value by $|y|$ and $|y|e^{-|y|^2}$ is integrable; therefore the dominated convergence theorem applies. We then obtain

$$g'(x) = -i\int_{\mathbb{R}} ye^{-|y|^2}e^{-ixy}dy.$$

By integration by parts this is equal to

$$-x\int_{\mathbb{R}} e^{-|y|^2}e^{-ixy} = -\frac{1}{2}xg(x).$$

Solving the differential equation $g'(x) = -\frac{1}{2}xg(x)$, we obtain

$$g(x) = ce^{-|x|^2/4}. \tag{16.7}$$

Also, we have $g(0) = \int_{\mathbb{R}} e^{-|y|^2}dy = \sqrt{\pi}$, so $c = \sqrt{\pi}$. Substituting this value of c in (16.7) and dividing both sides by $\sqrt{\pi}$ we obtain (16.6) in the case that $n = 1$.

If $n > 1$, we use Fubini's theorem to get

$$\begin{aligned}\hat{f}(x) &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}} e^{-|y_1|^2}e^{-ix_1y_1}dy_1\cdots\int_{\mathbb{R}} e^{-|y_n|^2}e^{-ix_ny_n}dy_n \\ &= \frac{1}{(2\pi)^n}\sqrt{\pi}e^{-|x_1|^2/4}\cdots\sqrt{\pi}e^{-|x_n|^2/4} \\ &= \frac{1}{(2\sqrt{\pi})^n}e^{-|x|^2/4}.\end{aligned}$$

□

Taking $f(x) = e^{-|x|^2}, \lambda = 1/\sqrt{2}$ in the formula

$$\widehat{\delta_\lambda f}(x) = \frac{1}{|\lambda|^n}\hat{f}\left(\frac{x}{\lambda}\right),\ x \in \mathbb{R}^n,$$

we get

$$\widehat{e^{-\frac{|\cdot|^2}{2}}}(x) = \frac{1}{(2\pi)^{n/2}}e^{-\frac{|x|^2}{2}}.$$

Consider the kernel

$$K(x) = \pi^{-n/2}e^{-|x|^2},\ x \in \mathbb{R}^n$$

and the corresponding approximation of the identity:

$$K_\epsilon(x) = \epsilon^{-n}K(x/\epsilon) = (\sqrt{\pi}\epsilon)^{-n}e^{-|x|^2/\epsilon^2},\ \epsilon > 0.$$

Note that $\int_{\mathbb{R}^n} K(x)dx = 1$. Setting $\epsilon = \sqrt{t}, t > 0$, we obtain the n-dimensional Gauss–Weierstrass kernel defined by

$$W(x,t) = (\sqrt{\pi t})^{-n}e^{-|x|^2/t},\quad x \in \mathbb{R}^n, t > 0. \tag{16.8}$$

Then

$$\int_{\mathbb{R}^n} W(x,t)dx = 1,\ t > 0.$$

For $x \in \mathbb{R}^n$ and $t > 0$, the convolution $Wf(x,t)$ defined by

$$Wf(x,t) = [f * W(\cdot,t)](x) = \int_{\mathbb{R}^n} f(x-y)W(y,t)dy \tag{16.9}$$

is called the Gauss–Weierstrass integral of f. The kernel $W(x,t)$ satisfies the heat equation

$$\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right)W(x,t) = 4\frac{\partial}{\partial t}W(x,t),\ x \in \mathbb{R}^n, t > 0.$$

Due to the dilation property and (16.6), the Gauss–Weierstrass kernel satisfies

$$W(x,t) = \widehat{e^{-\frac{t|\cdot|^2}{4}}}(x),\ \widehat{W(\cdot,t)}(x) = (2\pi)^{-n}e^{-t|x|^2/4}. \tag{16.10}$$

(10) (Shifting hats) If $f, g \in L^1(\mathbb{R}^n)$, then

$$\int_{\mathbb{R}^n} \hat{f}(x)g(x)dx = \int_{\mathbb{R}^n} \hat{g}(x)f(x)dx. \tag{16.11}$$

Proof. Since $f, g \in L^1(\mathbb{R}^n)$, $\hat{f}, \hat{g} \in L^\infty(\mathbb{R}^n)$, the integrals in (16.11) are finite. The formula is a corollary of Fubini's theorem since

$$\begin{aligned} \int_{\mathbb{R}^n} \hat{f}(x)g(x)dx &= \int_{\mathbb{R}^n} \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(y)e^{-ix\cdot y}dy \right) g(x)dx \qquad (16.12) \\ &= \int_{\mathbb{R}^n} f(y) \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} g(x)e^{-ix\cdot y}dx \right) dy \\ &= \int_{\mathbb{R}^n} \hat{g}(y)f(y)dy, \end{aligned}$$

where the change in the order of integration is justified because $f(y)g(x) \in L^1(\mathbb{R}^{2n})$. $\square$

Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be infinitely differentiable with compact support, non-negative, and with integral equal to 1. For $\epsilon > 0$, let

$$\phi_\epsilon(x) = \epsilon^{-n}\phi(x/\epsilon).$$

Theorem 16.2 *Suppose $1 \le p \le \infty$ and $f \in L^p(\mathbb{R}^n)$.*

(1) For each $\epsilon > 0$, $f * \phi_\epsilon$ is infinitely differentiable. For each $\alpha_1, \alpha_2, ..., \alpha_n$ non-negative integers

$$\frac{\partial^{\alpha_1+\cdots\alpha_n}(f*\phi_\epsilon)}{\partial_{x_1}^{\alpha_1}\cdots\partial_{x_n}^{\alpha_n}} = f * \frac{\partial^{\alpha_1+\cdots\alpha_n}(\phi_\epsilon)}{\partial_{x_1}^{\alpha_1}\cdots\partial_{x_n}^{\alpha_n}}$$

We use the convention that the 0^{th} order derivative of a function is just the function itself.

(2) $f * \phi_\epsilon \to f$ a.e. as $\epsilon \to 0$.

(3) If f is continuous, then $f * \phi_\epsilon \to f$ uniformly on compact sets as $\epsilon \to 0$.

(4) If $1 \le p < \infty$ and $f \in L^p(\mathbb{R}^n)$, then $f * \phi_\epsilon \to f$ in L^p.

Proof. (1) Suppose ϕ has support in $B(0, R)$. Let e_i be the unit vector in the i^{th} direction. Write

$$\begin{aligned} & f * \phi_\epsilon(x + he_i) - f * \phi_\epsilon(x) = \qquad (16.13) \\ = & \int_{\mathbb{R}^n} f(y)(\phi_\epsilon(x + he_i - y) - \phi_\epsilon(x - y))dy. \end{aligned}$$

Since ϕ_ϵ is continuously differentiable with compact support, there exists a constant c_1 such that

$$|\phi_\epsilon(x + he_i - y) - \phi_\epsilon(x - y)| \le c_1|h|$$

for all x and y. We may then divide both sides of (16.13) by h, let $h \to 0$, and use dominated convergence with the dominating function being $c_1|f(y)|\chi_{B(x,R\epsilon+1)}(y)$ provided $|h| \le 1$. This dominating function is in L^1 because $f \in L^p$.

We conclude
$$\frac{\partial(f*\phi_\epsilon)}{\partial x_i}(x) = f*\frac{\partial\phi_\epsilon}{\partial x_i}(x).$$
Since $\partial\phi_\epsilon/\partial x_i$ is also infinitely differentiable with compact support, we may continue and handle the higher order partial derivatives by induction.

(2) Since ϕ has integral 1, using a change of variables shows that ϕ_ϵ does too. We then have
$$\begin{aligned} f*\phi_\epsilon(x) - f(x) &= \int_{\mathbb{R}^n}(f(y)-f(x))\phi_\epsilon(x-y)dy \qquad (16.14)\\ &= \frac{1}{\epsilon^n}\int_{B(x,R\epsilon)}(f(y)-f(x))\phi(\frac{x-y}{\epsilon})dy. \end{aligned}$$
This leads to
$$\begin{aligned} |f*\phi_\epsilon(x)-f(x)| &\le \frac{1}{\epsilon^n}\int_{B(x,R\epsilon)}|f(y)-f(x)|\phi(\frac{x-y}{\epsilon})dy\\ &\le ||\phi||_\infty|B(0,R)|\frac{1}{|B(x,\epsilon R)|}\int_{B(x,R\epsilon)}|f(y)-f(x)|dy\\ . \end{aligned} \qquad (16.15)$$
The last line tends to 0 for almost every x by Theorem 13.7.

(3) Let $N>0$. Starting with (16.15),
$$\begin{aligned} &\sup_{|x|\le N}|f*\phi_\epsilon(x)-f(x)|\\ \le\ & ||\phi||_\infty|B(0,R)|\sup_{|x|\le N,|y-x|\le R\epsilon}|f(y)-f(x)|. \end{aligned}$$
This tends to 0 as $\epsilon\to 0$ because f is uniformly continuous on $B(0,N+R)$.

(4) Let $\epsilon>0$. If $f\in L^p$, let us take N sufficiently large and continuous g with compact support in $B(0,N)$ so that $||f-g||_p<\epsilon$.

By part (3) we have pointwise convergence of $g*\phi_\epsilon$ to g as $\epsilon\to 0$. Since ϕ_ϵ has support in $B(0,R\epsilon)$ and g has support in $B(0,N)$, we know that $g*\phi_\epsilon$ has support in $B(0,N+R\epsilon)$ and so $|g*\phi_\epsilon|\le ||g||_\infty||\phi||_\infty|B(0,R)|\chi_{B(0,N+R)}$ for $\epsilon<1$. We may therefore use dominated convergence to obtain that $g*\phi_\epsilon\to g$ in L^p as $\epsilon\to 0$.

Now $||f-g||_p<\epsilon$ and
$$||f*\phi_\epsilon - g*\phi_\epsilon||_p \le ||f-g||_p||\phi_\epsilon||_1<\epsilon,$$
so
$$||f*\phi_\epsilon - f||_p\le 2\epsilon+||g*\phi_\epsilon-g||_p.$$
Therefore
$$\lim_{\epsilon\to 0}||f*\phi_\epsilon - f||_p = 0$$
which proves (4). □

Proposition 16.3 *Suppose $\phi\in L^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n}\phi(x)dx=1$. Let $\phi_\delta(x)=\delta^{-n}\phi(x/\delta)$.*

(1) If g is continuous with compact support, then $g*\phi_\delta$ converges to g pointwise as $\delta\to 0$.

(2) If g is continuous with compact support, then $g*\phi_\delta$ converges to g in L^1 as $\delta\to 0$.

(3) If $f \in L^1$, then $||f * \phi_\delta - f||_1 \to 0$ as $\delta \to 0$.

Proof. (1) We have by a change of variables that $\int_{\mathbb{R}^n} \phi_\delta(x)dx = 1$. Then

$$\begin{aligned} |g * \phi_\delta(x) - g(x)| &= \left| \int_{\mathbb{R}^n} (g(x-y) - g(x))\phi_\delta(y)dy \right| \\ &= \left| \int_{\mathbb{R}^n} (g(x-\delta y) - g(x))\phi(y)dy \right| \\ &\leq \int_{\mathbb{R}^n} |g(x-\delta y) - g(x)||\phi(y)|dy. \end{aligned}$$

Since g is continuous with compact support and hence bounded and ϕ is integrable, the right hand side goes to zero by the dominated convergence theorem, the dominating function being $2||g||_\infty \phi$.

(2) We now use the Fubini theorem to write

$$\begin{aligned} \int_{\mathbb{R}^n} |g * \phi_\delta(x) - g(x)|dx &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (g(x-\delta y) - g(x))\phi(y)dy \right| dx \\ &\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x-\delta y) - g(x)||\phi(y)|dydx \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x-\delta y) - g(x)|dx|\phi(y)|dy. \end{aligned}$$

Let

$$G_\delta(y) = \int_{\mathbb{R}^n} |g(x-\delta y) - g(x)|dx.$$

By the dominated convergence theorem, for each $y, G_\delta(y)$ tends to 0 as $\delta \to 0$, since g is continuous with compact support. Moreover G_δ is bounded in absolute value by $2||g||_1$. Using the dominated convergence theorem again and the fact that ϕ is integrable, we see that $\int_{\mathbb{R}^n} G_\delta(y)|\phi(y)|dy$ tends to 0 as $\delta \to 0$.

(3) Let $\epsilon > 0$ and g be a continuous function with compact support so that $||f - g||_1 < \epsilon$. Let $h = f - g$. A change of variables shows that $||\phi_\delta||_1 = ||\phi||_1$. Observe

$$||f * \phi_\delta - f||_1 \leq ||g * \phi_\delta - g||_1 + ||h * \phi_\delta - h||_1.$$

Also

$$||h * \phi_\delta - h||_1 \leq ||h * \phi_\delta||_1 + ||h||_1 \leq ||h||_1||\phi_\delta||_1 + ||h||_1 < \epsilon(1 + ||\phi||_1)$$

by Theorem 15.3. Therefore, using (2),

$$\limsup_{\delta \to 0} ||f * \phi_\delta - f||_1 \leq \limsup_{\delta \to 0} ||h * \phi_\delta - h||_1 \leq \epsilon(1 + ||\phi||_1).$$

Since ϵ is arbitrary, we have our conclusion. □

(11) We have the following inversion result.

Theorem 16.4 (Inversion of the Fourier transform on L^1) *If f and $\hat{f}$ are in $L^1(\mathbb{R}^n)$, then*

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(y)e^{ix \cdot y}dy, \text{ a.e.} \tag{16.16}$$

Proof. Let $H_a(x) = e^{-\frac{|x|^2}{2a^2}}$; then we have from (16.10) that

$$\hat{H}_a(u) = W\left(u, \frac{2}{a^2}\right) = \frac{1}{(2\pi)^{n/2}} a^n e^{-\frac{a^2|u|^2}{2}}.$$

We write

$$\begin{aligned}\int_{\mathbb{R}^n} \hat{f}(u) e^{iu\cdot y} H_a(u) du &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-iu\cdot x} f(x) dx\, e^{iu\cdot y} H_a(u) du \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-iu(x-y)} H_a(u) du f(x) dx \\ &= \int_{\mathbb{R}^n} \hat{H}_a(x-y) f(x) dx \qquad (16.17)\end{aligned}$$

We can interchange the order of integration because

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x)||H_a(u)| dx du < \infty.$$

The left hand side of the first line of (16.17) converges to

$$\int_{\mathbb{R}^n} \hat{f}(u) e^{iu\cdot y} du.$$

as $a \to \infty$ by the dominated convergence theorem since $H_a(u) \to 1$ and $\hat{f} \in L^1$. The last line of (16.17) is equal to

$$\int_{\mathbb{R}^n} \hat{H}_a(y-x) f(x) dx = (f * \hat{H}_a)(y). \qquad (16.18)$$

But by Proposition 16.3, setting

$$\delta = a^{-1}, \ \phi(x) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{|x|^2}{2}},$$

we see that $f * \hat{H}_a = f * \phi_\delta$ converges to f in L^1 as $a \to \infty$. □

Corollary 16.5 (Uniqueness) *If $f, g \in L^1(\mathbb{R}^n)$ and $\hat{f}(x) = \hat{g}(x)$ for all $x \in \mathbb{R}^n$, then $f = g$ a.e. in $\mathbb{R}^n$.*

(12) (The Convolution Property) The convolution of any two integrable functions is also integrable and therefore has a well-defined Fourier transform.

Theorem 16.6 *If $f, g \in L^1(\mathbb{R}^n)$, then*

$$f \hat{*} g(x) = (2\pi)^n \hat{f}(x) \hat{g}(x) \ for\ all\ x \in \mathbb{R}^n.$$

This follows easily from Fubini's theorem and is left as an exercise.

(13) *(Differentiation Properties) We list two important differentiation formulas of the Fourier transform.*

Theorem 16.7 *(a) Let $x_k, k = 1, ..., n$, denote the kth coordinate of $x \in \mathbb{R}^n$. If f and $x_k f(x)$ belong to $L^1(\mathbb{R}^n)$, then $\hat{f}$ has a partial derivative with respect to*

x_k everywhere in $\mathbb{R}^n$, and

$$\frac{\partial \hat{f}}{\partial x_k} = (-i\hat{x_k f})\left[= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}(-iy_k f(y))e^{-ix\cdot y}dy\right] \tag{16.19}$$

(b) Fix $k = 1, ..., n$ and let $h = (0, ..., 0, h_k, 0, ..., 0) \neq 0$ lie on the kth coordinate axis. Suppose that $f \in L^1(\mathbb{R}^n)$ and that there is a function g such that

$$\lim_{h_k \to 0}\int_{\mathbb{R}^n}\left|\frac{f(x+h)-f(x)}{h_k} - g(x)\right| dx = 0. \tag{16.20}$$

Then $g \in L^1(\mathbb{R}^n)$ and

$$\hat{g}(x) = ix_k\hat{f}(x) \text{ for all } x \in \mathbb{R}^n. \tag{16.21}$$

Proof. (a) Denoting nonzero points on the kth coordinate axis by $h = (0, ..., 0, h_k, 0, ..., 0)$, we have

$$\frac{\hat{f}(x+h)-\hat{f}(x)}{h_k} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(y)\left\{\frac{e^{-ih_k y_k}-1}{h_k}\right\}e^{-ix\cdot y}dy.$$

Since the expression in curly brackets converges to $-iy_k$ as $h_k \to 0$ and is bounded in absolute value by $|y_k|$, formula (16.19) follows from the Lebesgue dominated convergence theorem and the hypothesis that $f(y)y_k$ is integrable.

(b) If g satisfies (16.20), then g is integrable since f is which implies that $\frac{f(x+h)-f(x)}{h_k} \in L^1(\mathbb{R}^n)$. With $h = (0, ..., 0, h_k, 0, ..., 0)$ as usual, (16.20) together with (16.2) implies the pointwise equality

$$\lim_{h_k \to 0}\frac{(f(\cdot + \hat{h}) - f(\cdot))}{h_k}(x) = \hat{g}(x).$$

Equivalently,

$$\lim_{h_k \to 0}\hat{f}(x)\left(\frac{e^{ih_k x_k}-1}{h_k}\right) = \hat{g}(x)$$

and (16.21) follows. □

A simple case when (16.20) holds with g equal to the ordinary partial derivative $\partial f/\partial x_k$ of f is when $f \in C_0^1(\mathbb{R}^n)$, since then both f and $\partial f/\partial x_k$ have compact support and the difference quotient of f in the variable x_k converges uniformly to $\partial f/\partial x_k$. Hence, by (16.21), for every $k = 1, ..., n$, we have

$$\hat{\frac{\partial f}{\partial x_k}}(x) = ix_k\hat{f}(x), \; x \in \mathbb{R}^n.$$

16.2. The Fourier Transform on L^2. We will now define the Fourier transform of a general $f \in L^2(\mathbb{R}^n)$ and study its main properties. We will see that a striking and fundamental difference between the maps $f \to \hat{f}$ when $f \in L^1(\mathbb{R}^n)$ and when $f \in L^2(\mathbb{R}^n)$ is that the mapping for $L^2(\mathbb{R}^n)$ turns out to be essentially an isometry of $L^2(\mathbb{R}^n)$ onto itself. Proving this is the main goal of the subsection.

Lemma 16.8 *Let $f, g \in L^2(\mathbb{R}^n)$. Then $f * g$ is uniformly continuous and bounded on $\mathbb{R}^n$, and $||f * g||_\infty \leq ||f||_2||g||_2$.*

Proof. Let $f, g \in L^2(\mathbb{R}^n)$. For any $x \in \mathbb{R}^n$,

$$\begin{aligned} |(f*g)(x)| &\le \int_{\mathbb{R}^n} |f(x-y)||g(y)|dy \quad (16.22)\\ &\le \left(\int_{\mathbb{R}^n} |f(x-y)|^2 dy\right)^{1/2} \left(\int_{\mathbb{R}^n} |g(y)|^2 dy\right)^{1/2} = ||f||_2||g||_2. \end{aligned}$$

In particular, $f*g$ is finite everywhere and $||f*g||_\infty \le ||f||_2||g||_2$. Also, if $x, h \in \mathbb{R}^n$, then

$$\begin{aligned} |(f*g)(x+h) - (f*g)(x)| &= \left|\int_{\mathbb{R}^n} [f(x+h-y) - f(x-y)]g(y)dy\right| \\ &\le ||f(\cdot + h) - f(\cdot)||_2||g||_2 \to 0 \text{ as } |h| \to 0 \end{aligned}$$

by continuity in L^2, and the proof is complete. □

We can now derive the key fact needed in order to extend the definition of the Fourier transform to $L^2(\mathbb{R}^n)$.

Theorem 16.9 *If f is continuous with compact support, then $\hat{f} \in L^2(\mathbb{R}^n)$ and*

$$||\hat{f}||_2 = (2\pi)^{-n/2}||f||_2. \quad (16.23)$$

Proof. First note that if we combine (16.17) and (16.18), then

$$\int_{\mathbb{R}^n} \hat{f}(u)e^{iu\cdot y}H_a(u)du = \int_{\mathbb{R}^n} \hat{H}_a(y-x)f(x)dx = (f*\hat{H}_a)(y).$$

Now take $y = 0$ to obtain

$$\int_{\mathbb{R}^n} \hat{f}(u)H_a(u)du = (f*\hat{H}_a)(0). \quad (16.24)$$

Set $g(x) = \overline{f(-x)}$; then $\hat{g}(x) = \overline{\hat{f}(x)}$. By (16.24) with f replaced by $f*g$,

$$\int_{\mathbb{R}^n} f\hat{*}g(u)H_a(u)du = (f*g*\hat{H}_a)(0). \quad (16.25)$$

Observe that

$$(f\hat{*}g)(u) = (2\pi)^n\hat{f}(u)\hat{g}(u) = (2\pi)^n|\hat{f}(u)|^2.$$

Thus the left hand side of (16.25) converges by the monotone convergence theorem to

$$(2\pi)^n \int_{\mathbb{R}^n} |\hat{f}(u)|^2 du$$

as $a \to \infty$. Since f and g are continuous with compact support, then $f*g$ is also, and so the right hand side of (16.25) converges to $f*g(0) = \int_{\mathbb{R}^n} f(y)g(-y)dy = \int_{\mathbb{R}^n} |f(y)|^2 dy$ by Proposition 16.3. □

Remark 16.10 We can use Theorem 16.8 to define $\hat{f}$ when $f \in L^2$ so that (16.23) will continue to hold. The set of continuous functions with compact support is dense in L^2 by Theorem 11.11. Given a function f in L^2, choose a sequence $\{f_m\}$ of continuous functions with compact support such that $f_m \to f$ in L^2. Then $||f_m - f_n||_2 \to 0$ as $m, n \to \infty$. By (16.23), $\{\hat{f_m}\}$ is a Cauchy sequence in L^2, and therefore converges to a function in L^2, which we call $\hat{f}$. Let us check that the limit does not depend on the choice of the sequence. If $\{f'_m\}$ is another sequence of continuous functions with compact support converging to f in L^2, then $\{f_m - f'_m\}$ is a sequence of continuous functions with compact support converging to 0 in L^2. By (16.23), $\{\hat{f_m} - \hat{f'_m}\}$ converges to 0 in L^2, and therefore $\hat{f'_m}$ has the same limit as $\hat{f_m}$. Thus $\hat{f}$ is defined uniquely up to almost everywhere equivalence. By passing to the limit in L^2 on both sides of (16.23), we see that (16.23) holds for $f \in L^2$.

Exercises 15

1. Prove Theorem 16.6.

2. Suppose that f and g are continuous with compact support. Show that $f * g$ is also continuous with compact support.

3. Let L be a nonsingular linear transformation of $\mathbb{R}^n$ and define $(Lf)(x) = f(Lx)$. If $f \in L^1(\mathbb{R}^n)$, show that

$$\hat{Lf}(x) = |detL|^{-1}\hat{f}(L^{-1})^*(x), \ x \in \mathbb{R}^n,$$

where $(L^{-1})^*$ is the adjoint of L^{-1}.

4. Let

$$W(x,t) = (\sqrt{\pi t})^{-n} e^{-|x|^2/t}, \quad x \in \mathbb{R}^n, t > 0. \tag{16.26}$$

Show that

$$\left(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right) W(x,t) = 4\frac{\partial}{\partial t} W(x,t).$$

17. Li-Yau inequality

In this section, we introduce the famous Li-Yau inequality about eigenvalues of the Laplacian with Dirichet boundary condition which is a very beautiful application of the Fourier transformation.

Let Ω be a bounded domain in $\mathbb{R}^n$ and Δ the Laplace operator of $\mathbb{R}^n$. Consider the Dirichlet eigenvalue problem:

$$\left\{ \begin{array}{rcl} \Delta u &=& -\lambda u, \quad \text{in} \quad \Omega, \\ u|_{\partial\Omega} &=& 0, \end{array} \right. \tag{17.1}$$

where u is nonzero function and λ is a real number. We call u an eigenfunction corresponding to the eigenvalue λ. The eigenvalues of this problem can be arranged as

$$0 < \lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \to \infty.$$

In 1912, Weyl proved the following asymptotic formula

$$\lambda_k \sim C(n)\left(\frac{k}{|\Omega|}\right)^{2/n} \quad \text{as } k \to \infty \tag{17.2}$$

where $C(n) = (2\pi)^2 B_n^{-2/n}$, B_n is the volume of the unit ball in $\mathbb{R}^n$. Based on Weyl's formula, Pólya proposed the following conjecture:

$$\lambda_k \geq C(n)\left(\frac{k}{|\Omega|}\right)^{2/n} \quad \forall k = 1, 2, \cdots . \tag{17.3}$$

An important progress to Pólya's conjecture is the Li-Yau inequality.

Theorem 17.1 (Li-Yau, 1983)

$$\sum_{i=1}^{k} \lambda_i \geq \frac{nkC(n)}{n+2}\left(\frac{k}{|\Omega|}\right)^{2/n}. \tag{17.4}$$

In particular,

$$\lambda_k \geq \frac{nC(n)}{n+2}\left(\frac{k}{|\Omega|}\right)^{2/n}. \tag{17.5}$$

Before proving Theorem 17.1, we prove a

Lemma 17.2 *If f is a real-valued function defined on $\mathbb{R}^n$ with $0 \leq f \leq M_1$ and*

$$\int_{\mathbb{R}^n} |z|^2 f(z)dz \leq M_2,$$

then

$$\int_{\mathbb{R}^n} f(z)dz \leq (M_1 B_n)^{\frac{2}{n+2}} M_2^{\frac{n}{n+2}} \left(\frac{n+2}{n}\right)^{\frac{n}{n+2}}.$$

Proof. Let $g(z) = M_1$ when $|z| < R$ and $g(z) = 0$ when $|z| > R$ where R is chosen so that

$$\int_{\mathbb{R}^n} |z|^2 g(z)dz = M_2.$$

Then $(|z|^2 - R^2)(f(z) - g(z)) \geq 0$, hence

$$R^2 \int_{\mathbb{R}^n} (f(z) - g(z))dz \leq \int_{\mathbb{R}^n} |z|^2 (f(z) - g(z))dz = 0.$$

We have

$$M_2 = \int_{\mathbb{R}^n} |z|^2 g(z)dz = M_1 \int_0^R r^{n+1}\omega_{n-1}dr = M_1\left(\frac{\omega_{n-1}}{n+2}\right)R^{n+2}$$

where ω_{n-1} is the volume of the unit $(n-1)$ sphere in $\mathbb{R}^n$, and $\int_{\mathbb{R}^n} g(z)dz = M_1 B_n R^n$. Hence using $nB_n = \omega_{n-1}$ we conclude that

$$\int_{\mathbb{R}^n} f(z)dz \leq \int_{\mathbb{R}^n} g(z)dz = (M_1 B_n)^{\frac{2}{n+2}} M_2^{\frac{n}{n+2}} \left(\frac{n+2}{n}\right)^{\frac{n}{n+2}}.$$

Proof of Theorem 17.1. Let $\{u_i\}_{i=1}^k$ be the set of orthonormal eigenfunctions for the eigenvalues $\{\lambda_i\}_{i=1}^k$. That is,

$$\begin{cases} \Delta u_i &= -\lambda_i u_i, \quad \text{in} \quad \Omega, \\ u|_{\partial\Omega} &= 0, \\ \int_\Omega u_i u_j dx &= \delta_{ij}. \end{cases} \tag{17.6}$$

We consider the function defined by

$$F(x,y)=\sum_{i=1}^{k}u_i(x)u_i(y) \tag{17.7}$$

for $x,y\in\Omega$. The normalized Fourier transform of F in the x-variable is given by

$$\tilde{F}(z,y)=(2\pi)^{-n/2}\int_{\mathbb{R}^n}F(x,y)e^{-x\cdot z}dx. \tag{17.8}$$

It has the property that

$$\int_{\mathbb{R}^n}|\tilde{F}|^2(z,y)dz=\int_{\mathbb{R}^n}F^2(x,y)dx. \tag{17.9}$$

Hence

$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|\tilde{F}|^2(z,y)dzdy &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}F^2(x,y)dxdy \\ &= \int_{\Omega}\int_{\Omega}F^2(x,y)dxdy=k\end{aligned} \tag{17.10}$$

by the orthonormality of $\{u_k\}$.

On the other hand,

$$\begin{aligned}\int_{\mathbb{R}^n}|\tilde{F}|^2(z,y)dy &= \int_{\Omega}(2\pi)^{-n}\left|\int_{\mathbb{R}^n}F(x,y)e^{-ix\cdot y}dy\right|^2dx \\ &= \int_{\Omega}(2\pi)^{-n}\left|\int_{\Omega}F(x,y)e^{-ix\cdot y}dy\right|^2dx.\end{aligned} \tag{17.11}$$

By the definition of $F(x,y)$, this is a multiple by $(2\pi)^{-n}$ of the square of the L^2-norm of the projection of the function $h(x)=e^{-ix\cdot z}$ onto the subspace spanned by the first k eigenfunctions. Hence

$$\int_{\mathbb{R}^n}|\tilde{F}|^2(z,y)dy\le(2\pi)^{-n}\int_{\Omega}|e^{-ix\cdot z}|^2dx=(2\pi)^{-n}|\Omega| \tag{17.12}$$

which is the square of the L^2-norm of $h(x)$.

We consider the equalities

$$\begin{aligned}z_j\tilde{F}(z,y) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n}F(x,y)z_je^{-ix\cdot z}dx \\ &= (2\pi)^{-n/2}\int_{\mathbb{R}^n}F(x,y)i\frac{\partial e^{-ix\cdot z}}{\partial x_j}dx \\ &= (-i)(2\pi)^{-n/2}\int_{\mathbb{R}^n}\frac{\partial F(x,y)}{\partial x_j}e^{-ix\cdot z}dx \\ &= (-i)\frac{\partial\tilde{F}}{\partial x_j}(z,y)\end{aligned} \tag{17.13}$$

which implies that

$$\begin{aligned}
\int_{\mathbb{R}^n}\int_\Omega |z|^2|\tilde{F}|^2(z,y)dydz &= \int_{\mathbb{R}^n}\int_\Omega |\tilde{\nabla_x F}|^2(x,y)dydx \\
&= \int_{\mathbb{R}^n}\int_\Omega |\nabla_x F|^2(x,y)dydx \\
&= -\int_\Omega\int_\Omega F(x,y)\Delta_x F(x,y)dydx \\
&= \sum_{i=1}^k \lambda_i \qquad (17.14)
\end{aligned}$$

by definition of F.

Now, we can apply Lemma 17.2 to the function

$$f(z) = \int_{\mathbb{R}^n} |\tilde{F}|^2(z,y)dy \qquad (17.15)$$

with $M_1 = (2\pi)^{-n}|\Omega|$ and $M_2 = \sum_{i=1}^k \lambda_i$. We conclude that

$$k = \int_{\mathbb{R}^n} f(z)dz \le ((2\pi)^{-n}|\Omega|B_n)^{\frac{2}{n+2}}\left(\frac{n+2}{n}\right)^{\frac{n}{n+2}}\left(\sum_{i=1}^k \lambda_i\right)^{\frac{n}{n+2}}.$$

Hence,

$$\sum_{i=1}^k \lambda_i \ge \frac{nkC(n)}{n+2}\left(\frac{k}{|\Omega|}\right)^{2/n}.$$

$\square$